Structural Characterisation of Fungal and Bacterial Amyloids by Hydrogen/Deuterium Exchange NMR Spectroscopy

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades einer

Doktorin der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Agnes Zimmer

aus Weißenfels

1. Referentin: Prof. Dr. Christiane Ritter
2. Referent: Prof. Dr. Michael Steinert
eingereicht am: 15.08.2011
mündliche Prüfung (Disputation) am: 05.12.2011

Druckjahr 2012

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch die Mentorin der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen:

Wasmer, C., Zimmer, A., Sabaté, R., Soragni, A., Saupe, S. J., Ritter, C., and Meier, B. H. (2010). Structural similarity between the prion domain of HET-s and a homologue can explain amyloid cross-seeding in spite of limited sequence identity. J Mol Biol *402,* 311-325

Tagungsbeiträge:

Ritter, C., Nagaraj, M., Zimmer, A.,: NMR Structural investigations of the curli subunits CsgA and CsgB (Poster) FASEB Summer Research Conference - The Basic Origins and Medical Consequences of Protein Aggregation, Snowmass Village, USA (2011)

Zimmer, A., Eberth A., Lührs T., Ritter C.: Structural investigations of the major curlin subunit CsgA – a functional amyloid (Poster) 18th International MPSA conference, Uppsala, Sweden (2010)

Zimmer, A., Benkemoun, L., Wasmer, C., Meier, B.H., Saupe, S.J., Ritter, C.: Hydrogen exchange studies of HET-s homologue capable for species barrier breaching (Poster) Benzon Symposium 56 – Functioanal and Pathogenic Protein Aggregation, Copenhagen, Denmark (2009)

Zimmer, A., Wasmer, C., Saupe, S.J., Meier, B.H., Ritter C.: Hydrogen exchange studies of a HET-s homolog from *Fusarium graminearum* (Poster) International Bunsen Discussion Meeting on Structure of Amyloid Fibrils and Mechanism of Amyloid Formation, Halle/Saale (2009)

Zimmer, A., Lührs, T., Ritter, C.: Characterization of amyloid fibrils by quenched hydrogen exchange NMR. (Vortrag) 1st International Meeting “Conformational Transitions in Macromolecular Interactions”, Halle/Saale (2008)

Des Nutzens wegen ist die Natur gezwungen,
so zu verfahren, wie sie es tut.

Mark Aurel, Selbstbetrachtungen

Contents

Abbreviations

A_x	Absorption (x is any wavelength)
aa	Amino acid
Å	Ångström
AFM	Atomic force microscopy
AF_{488}	AlexaFluor488
AP	Alkaline phosphatise
App	Appendix
APS	Ammonium persulphate
BMRB	Biological magnetic resonance data bank
BCIP	5-bromo-4-chloro-3-indolylphosphate
Bis-Tris	Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methan
CD	Circular dichroism
CR	Congo red
CsgX	Genproduct of curli specific gene X
CV	Column volume
Da	Dalton
E. coli	*Escherichia coli*
Ec	*Escherichia coli*
EM	Electron microscopy
ECL	Electrochemiluminescence
DMSO	Dimethyl sulfoxide
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
FA	Formic acid
Fg	*Fusarium gramineaum*
H/D	Hydrogen/Deuterium
hNOE	Heteronuclear Overhauser effect
HSQC	Heteronuclear single quantum coherence
FT-IR	Fourier transform infrared
HRP	Horse radish peroxidase
IAA	Iodoacetic acid

IAM	Iodoacetamide
IB	Inclusion body
IMAC	Immobilized metal ion affinity chromatography
IPTG	Isopropyl-β-D-thiolgalactoside
KPi	KH_2PO_4/K_2HPO_4
MHC	Major histocompatibility complex
MWCO	Molecular weight cut off
NBT	Nitro blue tetrazolium
Ni-NTA	Nickel-nitrilotriacetic acid-agarose
NOE	Nuclear Overhauser Effect
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser enhancement spectroscopy
OD_{600}	Optical density at 600 nm
Pa	*Podospora anserina*
PAGE	Polyacrylamid gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank
PM	N-(1-pyrene)maleimide
R	Repeat
SDS	Sodium dodecyl sulphate
Sec	Secretory
TAE	Tris-acetate-EDTA
TBS	Tris buffered saline
TCA	Trichloroacetic acid
TCEP	Tris(2-carboxyethyl)phosphine
TEM	Transmission electron microscopy
TEMED	N, N, N', N' – Tetramethylendiamin
TFA	Trifluoroacetic acid
ThT	Thioflavin T
Tris	Tris(hydroxymethyl)aminomethane
wt	Wild type

Amino acids are abbreviated using 1 or 3 letter code.

Latinisms are written in italic.

Zusammenfasung

In den letzten Jahren wurde für eine steigende Anzahl von Amyloiden gezeigt, dass diese physiologisch relevante Zellfunktionen übernehmen. Im Gegensatz zu krankheitsassoziierten Amyloiden fibrillisieren die Proteine ohne toxischen Nebeneffekt in geordnete, β-faltblattreiche Strukturen. Das HET-s Prion aus *Podospora anserina* und die Curli-Fibrillen aus *E. coli* gehören zu den bisher bestuntersuchten funktionalen Amyloiden. Entsprechend ihrer Funktionen stellen sie einzigartige Systeme dar, um den Prozess der Fibrillenbildung und auch die strukturellen Grundlagen von Prionentransmission und die Rolle bakterieller Amyloide als Pathogenitätsfaktoren zu untersuchen.

Im Vorfeld wurde gezeigt, dass die Priondomäne des HET-s Homologen aus *Fusarium graminearum* *Fg*HET-$s_{218-289}$ *in vitro* Fibrillen bildet, die trotz geringer Sequenzidentität die Fibrillisierung von *Pa*HET-$s_{218-289}$ nukleieren. In der vorliegenden Arbeit wurde *Fg*HET-$s_{218-289}$ rekombinant hergestellt, gereinigt, fibrillisiert und mittels Wasserstoff/Deuterium Austauschexperimenten untersucht. Zusammen mit Festkörper-NMR Daten belegten die Austauschraten den Grund für den Nukleationseffekt: Es wurde gezeigt, dass die Proteine strukturell eng verwandt sind. Beide Priondomänen formen eine sehr ähnliche β-Solenoidstruktur, die die Nukleation ermöglicht. Der dreieckige hydrophobe Kern und wichtige Strukturelemente sind evolutiv konserviert.

Die Topologie des Hauptbestandteils der Curli-Fibrillen CsgA aus *E. coli* wurde erfolgreich mittels Austauschexperimenten und Thiol-Markierungsexperimenten untersucht. Diesbezüglich wurde die Herstellung der Fibrillen optimiert, zusätzliche NMR-Experimente, selektive Isotopenmarkierung sowie zahlreiche *Ec*CsgA Varianten wurden eingesetzt. Die nachfolgende Analyse zeigte, dass der Amyloidkern strukturell repetitiv aufgebaut ist. Es wurde vorgeschlagen, dass *Ec*CsgA eine β-Solenoidstruktur bildet. Zudem wurden heteronukleierte Fibrillen untersucht. Die Untersuchungen zeigten, dass *in vitro* generierte *Ec*CsgB Fibrillen kein geeigneter Nukleationskeim sind.

Um lösliche amyloid-ähnliche Faltungsintermediate von *Ec*CsgA zu untersuchen, wurde ein Set von Fusionsproteinen generiert und hinsichtlich Löslichkeit und amyloidogener Faltung untersucht. Es wurden zwei Analoge identifiziert, die löslich hergestellt werden können und teilweise stabil gefaltet sind

Die erzielten Ergebnisse stellen die ersten Sequenz-spezifischen Strukturdaten für *Ec*CsgA Fibrillen dar und bilden die Grundlage für weiterführende Untersuchungen möglicher Struktur-Funktionsbeziehungen.

Summary

In recent years, an increasing number of proteins has been shown to form amyloids that are involved in a large variety of cellular processes. Contrary to disease-related amyloids, these proteins fibrillise without toxic side effect into well ordered β-sheet rich structures. The HET-s prion of *Podospora anserina* and the curli system of *E. coli* are two of the most studied functional amyloids. Owing to their biological roles, these two amyloids constitute a group of unique systems that allows the analysis of amyloidogenesis as well as the structural background of prion transmission and the role of bacterial amyloids in pathogenicity.

The prion domain of a distinct homologue of HET-s from *Fusarium graminearum* *Fg*HET-s$_{218-289}$ was previously shown to fibrillise *in vitro* and to cross-seed *Pa*HET-s$_{218-289}$ prion formation despite low sequence identity. In the present study, *Fg*HET-s$_{218-289}$ was recombinantly produced, purified and fibrillised for hydrogen/deuterium exchange NMR experiments. In conjunction with solid state NMR experiments, obtained exchange data readily explained the HET-s cross-seeding phenomenon. The analysis demonstrated that the structures of both prion domains are closely related. Both prion domains constitute the same β-solenoid fold enabling the nucleation effect. The triangular hydrophobic core and major structural elements were identified to be conserved.

The topology of the major curli subunit CsgA of *E. coli* was further described by hydrogen/deuterium exchange analysis and thiol-labelling studies. The analysis comprised an optimised sample preparation, the implementation of different NMR experiments, new *Ec*CsgA constructs and selective isotope labelling. The approach revealed a repetitive structural pattern with five imperfect repeat sequences forming each a strand-turn-strand motif. Altogether the amyloid core of *Ec*CsgA was suggested to fold into a β-solenoid. Furthermore, heteronucleated *Ec*CsgA fibrils were also subjected to hydrogen/deuterium exchange experiments. The corresponding experiments revealed that *in vitro* formed *Ec*CsgB fibrils are no adequate seed.

In order to analyse amyloid-like intermediates of *Ec*CsgA, numerous fusion proteins were generated and analysed in consideration of their solubility and (non)amyloidogenic behaviour. At least two protein constructs were identified to be soluble and folded to a certain extend.

These results present the first sequence-specific data of structural elements of *Ec*CsgA fibrils and provide the background for following structure-function analyses.

1 Introduction

1.1 Amyloids and Prions – Friends or Foes?

1.1.1 Protein Folding

Proteins are essential components of the living organism and involved in nearly every process taking place in the cell. Thus, proteins are key players in the cell with regard to biochemical catalysis, regulation, transport and storage, immune response, maintaining the structure of other macromolecules, signalling processes and converting chemical into mechanical energy (Stryer, 1996). The native, three-dimensional structure of every protein dictates its function and is thus indispensable for the underlying biological processes.

The specific three-dimensional structure of every protein is encoded by its primary structure (Anfinsen, 1973). The polypeptide chain is folded through hydrogen bonds, ionic and hydrophobic interactions as well as covalent bonds (disulfide bridges) forming typical secondary structural elements and finally a characteristic tertiary structure. Oligomeric proteins assemble from more than one polypeptide chain and they feature a quaternary structure (Linderstrøm-Lang, 1952). Although protein folding has been studied intensely during the last decades, the correlation between primary structure and folding process still remains unclear. Nevertheless it is known that the folding pathway is driven thermodynamically to adopt the energetically most favourable conformation (Jaenicke, 1996). Considering the energy landscape scheme of protein folding and aggregation, misfolding and even alternative conformations are possible (Hartl and Hayer-Hartl, 2009; Herczenik and Gebbink, 2008). Thus, proteins are highly flexible structures and can adopt a large variety of conformations also leading to misfolding and aggregation (**Figure 1-1**) as well as enhancing functionality of a protein (Boehr and Wright, 2008; Murzin, 2008). Nevertheless, exposing hydrophobic amino acid residues and flexible regions, normally buried in the native fold, partially folded/unfolded and misfolded states are quite aggregation prone as aggregation is largely driven by hydrophobic forces. The aggregation pathway can result in amorphous/disordered aggregates or alternatively in highly ordered fibrillar aggregates called amyloid fibrils. These fibrils are characterised by the extreme robust cross-β structure (Sunde *et al.*, 1997). In the case of prions, this conformation is self-propagating and horizontally transmissible (Aguzzi *et al.*, 2008; Prusiner, 1998).

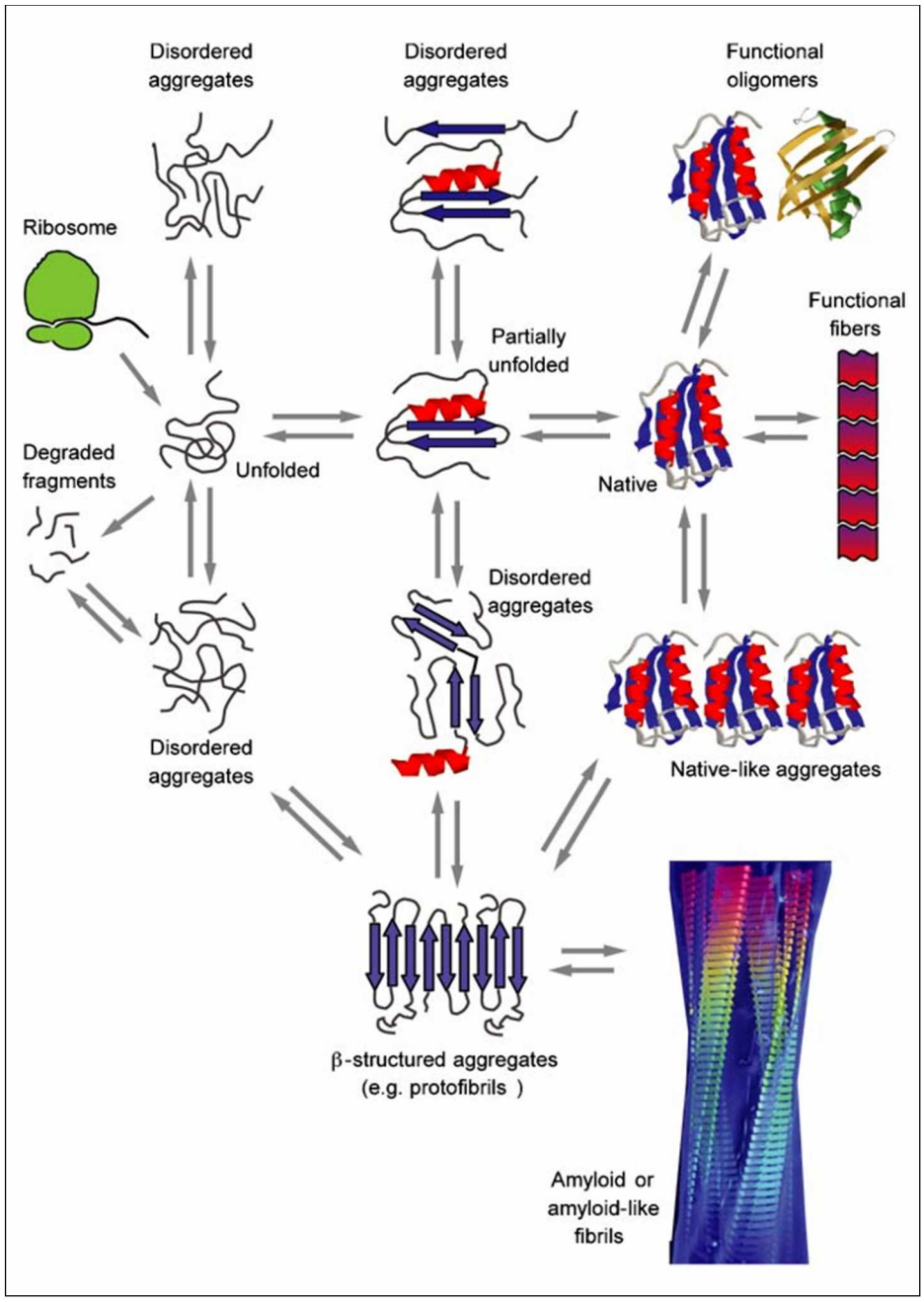

Figure 1-1: Some of the conformational states proteins can adopt. Proteins can adopt different conformations. The transition of protein conformations is highly regulated in the living cell by chaperons, degradation and protein quality control. Most conformations are functional but in case of failed regulation misfolding diseases occur (original figure from Chiti and Dobson, 2006).

In vivo, a certain degree of protein misfolding occurs regularly and is usually regulated and eliminated by the protein quality control via chaperones and degradation (Gregersen *et al.*, 2006). A number of misfolding events overwhelms the protein quality control and leads to protein misfolding diseases. These diseases are described by large deposits of aggregates, which can consist of amyloid or non-amyloid systems. An increasing number of proteins has been shown to form amyloid fibrils related to devastating amyloidoses like Alzheimer´s disease, type 2 diabetes, polyneuropathy syndromes and prion diseases

illustrated in **Table 1-1** (reviewed by Chiti and Dobson, 2006; Herczenik and Gebbink, 2008).

In contrast to disease related aggregates, more and more proteins have been shown to form amyloids as their native fold, which fulfil normal cellular functions summarised in **Table 1-2**. In bacteria these functions comprise for example biofilm formation, cell-cell recognition, host attachment and invasion, as well as sporulation, resistance and adaptation to environmental stress (reviewed by Chiti and Dobson, 2006; Gebbink *et al.*, 2005; Wickner *et al.*, 2007). At the beginning of the 21st century it was shown for the first time that proteins, which do not form amyloids *in vivo*, are able to form amyloid fibrils *in vitro* (Damaschun *et al.*, 2000; Fandrich *et al.*, 2001). These findings already suggested that amyloid formation is a generic property of polypeptides (Dobson, 2004).

1.1.2 Amyloid Fibrils Exhibit Unique Properties

Generally, amyloid formation is a poorly understood pathway comprising the accumulation of monomeric proteins into well ordered fibrillar aggregates. A nucleation dependent pathway is suggested in the Landsbury model (Rochet and Lansbury, 2000). The i*n vitro* polymerisation process is believed to pass through three distinct phases (Lomakin *et al.*, 1997). First, a stable β-sheet enriched nucleus is formed in the lag phase. Subsequently, in the growth phase the nucleus rapidly extends to higher oligomers and fibrils by self-assembly. The fibril formation is finally completed in the stationary phase. Characteristic of amyloid fibrils in this phase is the cross-β structure (**Figure 1-2 A**).

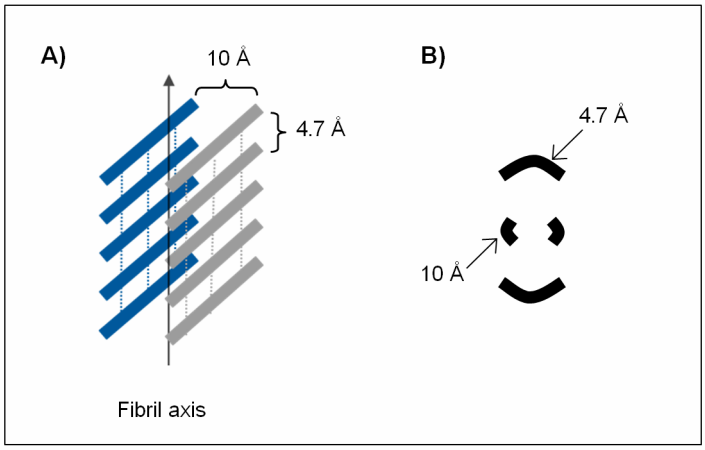

Figure 1-2: The amyloid cross-β structure. (A) The cross-β structure with β-strands perpendicular and β-sheets parallel to the main fibril axis (arrow). Dashed lines indicate hydrogen bonds. **(B)** Schematic fibril diffraction pattern that shows characteristic diffraction lines (adapted from Makin and Serpell, 2005).

X-ray fibril diffraction experiments identified β-strands perpendicular to the main fibril axis and β-sheets parallel to this axis. The two sets of diffraction lines of the obtained "cross" pattern (**Figure 1-2 B**) correspond to interstrand (~ 4.7 Å meridian) and stacking distances (~ 10 Å equator) in the β-sheets (Wormell, 1954; Sunde *et al.*, 1997).

The β-strands of a β-sheet can be arranged parallel or antiparallel to each other. In amyloid fibrils the parallel organisation is most common (Shewmaker *et al.*, 2009; Toyama and Weissman, 2011) and the β-strands are coordinated in register or in a β-solenoid. In parallel in register β-sheets identical β-strands are aligned on top of each other. Two or more β-sheets face each other forming a β-sandwich pattern which is continued by stacking interactions. The alternative motif is the β-solenoid, which is formed by repeating units of coils. Each coil consists of 2 - 4 β-strands and is elongated by an axial rise of 4.8 Å (Kajava and Steven, 2006). These coils are aligned on top of each other to form the corresponding β-sheets. β-solenoids are further distinguished in β-rolls and β-helices. Similar to the β-sandwich, β-rolls are made of two sheets forming an interface. The core of a β-helix consists of more than two β-sheets. An overview of the possible β-sheet architectures is given in **Figure 1-3** (adapted from Toyama and Weissman, 2011).

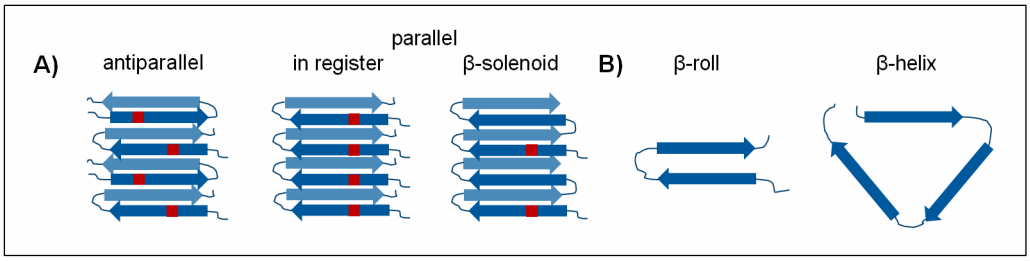

Figure 1-3: β-sheet architectures that proteins can adopt. (A) β-strands can be arranged antiparallel, parallel in-register or in a β-solenoid. These forms can be distinguished by the distances between single residues/probes (red rectangle, see text above). **(B)** Top view of the β-solenoid forms β-roll and β-helix (adapted from Toyama and Weissman, 2011).

Moreover, a β-sheet motif, called a "dry steric zipper", with an unusually high density was postulated as a basic principle of the cross-β structure (Nelson *et al.*, 2005; Sawaya *et al.*, 2007). These tight interfaces are mediated by van der Waals linkages and contain no water molecules (**Figure 1-4 A**).

Amyloids were identified as long, unbranched fibrous structures with a diameter of 5 - 15 nm and several micrometers in length (**Figure 1-4 B**) by transmission electron microscopy (TEM) and atomic force microscopy (AFM). The fibrils consist of several protofilaments that surround a main axis in a helical manner and interact with each other via amino acid side chains illustrated in **Figure 1-4 C** (Jimenez *et al.*, 1999; Jimenez *et al.*,

2002; Makin and Serpell, 2005; Zhang *et al.*, 2009). Amyloids are often characterised by structural polymorphism regarding the number of filaments per fibril, degree of twisting and the diameter or mass per fibril length (Fandrich *et al.*, 2009).

Furthermore, most amyloid fibrils can be visualised indirectly due to their affinity for amyloid specific dyes (Hawe *et al.*, 2008). In the presence of these specific aggregates, Thioflavin T (ThT) reveals a strong fluorescence emission around 480 – 490 nm and the binding of Congo Red results in a green birefringence in cross-polarised light (**Figure 1-4 D**).

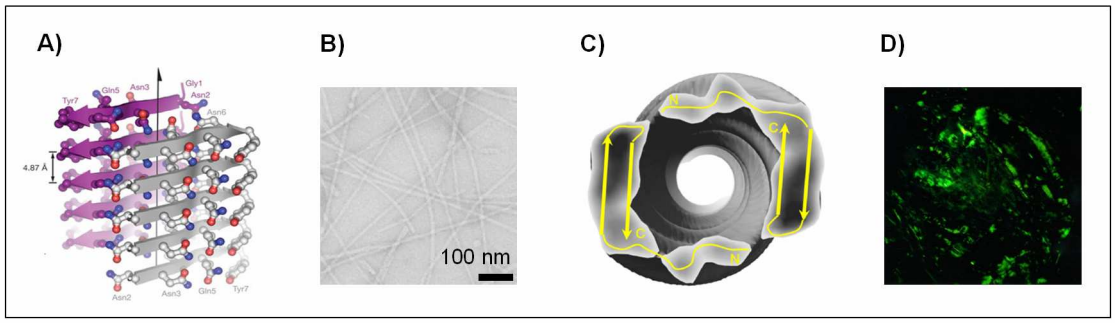

Figure 1-4: Properties of amyloid fibrils. (A) Atomistic structure of microcrystals of the peptide GNNQQNY indicating the "dry steric zipper" (Nelson *et al.*, 2005). Each β-strand represents one peptide molecule. **(B)** Transmission electron micrograph (TEM) of fibrils formed by islet amyloid polypeptide. (Makin and Serpell, 2005). **(C)** Model of interprotofilament interactions of $A\beta_{1-42}$ amyloid fibrils obtained by cryo electron microscopy (cryo EM) (Zhang *et al.*, 2009). The atomic model (pdb entry 2BEG) fits into the cryo EM density map. **(D)** Electron micrograph of amyloid deposits in presence of Congo Red in cross-polarised light (Thompson *et al.*, 2000).

In comparison to other proteins and unspecific aggregates, amyloids are extremely stable. Many fibrils are resistant against 1 % (w/v) SDS (Brown *et al.*, 1990; Collinson *et al.*, 1999) and are often only soluble in hexafluoroisopropanol, 90 % (v/v) formic acid or highly concentrated solutions of urea or guanidine hydrochloride (Chiba *et al.*, 2003; Collinson *et al.*, 1999; Nichols *et al.*, 2005; Wang *et al.*, 2007). Due to their compact structure, amyloids are protected against proteolytic degradation (Collinson *et al.*, 1999; Meyer *et al.*, 2000).

1.1.3 The Toxicity of Amyloid Fibril Formation

In the past, the formation of amyloid fibrils has only been associated with fatal protein misfolding diseases as the long list of amyloid proteins related to amyloidoses in **Table 1-1** shows (reviewed by Chiti and Dobson, 2006; Herczenik and Gebbink, 2008). Amyloidoses describe disease patterns of large abnormal deposits of amyloid fibrils in organs and/or

tissues featuring cell death and the loss of function of the affected organ/tissue. The dysfunction can be inherited or acquired, systemic or localised. Neurodegenerative diseases as for example Alzheimer's disease or the spongiform encephalopathies are amyloidoses of the nervous system and correspond to the progressive loss of structure and function of neurons. Amyloidoses, which do not affect the nervous system, are nonneuropathic diseases subdivided into systemic and localised amyloidoses. While localised amyloidoses are limited to only one organ or tissue, systemic amyloidoses correspond to amyloid deposits in more than one organ or system.

Table 1-1: Amyloidoses - human disease associated with amyloid fibril formation (reviewed by Chiti and Dobson, 2006; Herczenik and Gebbink, 2008).

Disease	Associated proteins	Native structure	Affected tissue
Neurodegenerative diseases			
Alzheimer´s disease	Amyloid β, tau	Unfolded	Brain
Spongiform encephalopathies	Prion protein	Unfolded N-term. α-helical C-term.	Brain, peripheral nervous system
Parkinson´s disease	α-Synuclein	Unfolded	Brain
Lewy-body dementia	α-Synuclein	Unfolded	Brain
Huntington´s disease	Huntingtin	Largely unfolded	Brain
Amyotrophic lateral sclerosis	Superoxide dismutase I	All-β, Ig like	Brain
Familiar amyloidic polyneuropathy	Transthyretin	All-β, prealbumin like	Peripheral nervous system
Nonneuropathic systemic amyloidoses			
Primary systemic amyloidosis	Immunoglobulin light chain	All-β, Ig like	Most tissues
Secondary (reactive) systemic amyloidosis	Serum amyloid A	All-α, unknown fold	Most tissues
Senile systemic amyloidosis	Transthyretin	All-β, prealbumin like	Microvasculature
Hemodialysis-related amyloidosis	β_2-Microglobulin	All-β, Ig like	Osteoarticular tissues
Nonneuropathic localised amyloidoses			
Type II diabetes	Amylin (islet amyloid polypeptide IAPP)	Unfolded	Pancreas (islet)
Injection localised amyloidosis	Insulin	All-α, insulin-like	Skin, muscles
Cataract	Crystallin	All-β, crystalline-like	Eye
Familial subepithelial corneal amyloidosis	Lactoferrin	α+ β, periplasmic-binding protein like	Cornea

Even though all diseases are associated with large amyloid deposits, there is a controversy about the pathogenic agent (Golde *et al.*, 2006). Initially, only the aggregated proteins/peptides were believed to be the toxic form of the protein leading to apoptosis

(amyloid hypothesis, Yankner *et al.*, 1990). Today, oligomeric and protofibril intermediates of the fibrillisation pathway as well as off-pathway oligomers are believed to be the toxic agent (Baskakov *et al.*, 2002; Hardy and Selkoe, 2002; Kayed *et al.*, 2003). These oligomers are reported to permeate the membrane and to form ion channels (Kourie and Henry, 2002; Lashuel, 2005). The membrane disruption is suggested to be a key player in the pathogenesis of amyloidoses (Lashuel, 2005; Quist *et al.*, 2005). It leads to an uncontrolled ion transport (especially Ca^{2+}), which disturbs the ion equilibrium between cytoplasm, mitochondria and endoplasmatic reticulum. Finally, the dysfunction in homeostasis causes apoptosis. The propensity to form ion-permeable pores via nonfibrillar oligomers has been shown for several amyloid forming proteins: Aβ (Hirakura *et al.*, 1999; Lashuel *et al.*, 2002a), amylin (Hirakura *et al.*, 2000; Mirzabekov *et al.*, 1996), serum amyloid A (Hirakura *et al.*, 2002), α-synuclein (Lashuel *et al.*, 2002b; Volles and Lansbury, 2002; Volles *et al.*, 2001) and even prion proteins (Kourie, 2002; Kourie *et al.*, 2001). Nevertheless, comprehension of initiation and progress of neurodegeneration and cytotoxicity still remains limited.

1.1.4 Amyloid Fibrils as Functional Fold

Recently, an increasing number of proteins has been shown to form amyloid fibrils, which are nontoxic and can be related to normal cellular functions presented in **Table 1-2** (reviewed by Chiti and Dobson, 2006; Fowler *et al.*, 2007; Gebbink *et al.*, 2005; Greenwald and Riek, 2010; Otzen and Nielsen, 2008). A wide range of organisms, comprising species from bacteria to mammals, takes advantage of the intrinsic propensity of proteins to form highly stable amyloid fibrils.

Table 1-2: Proteins that form functional amyloid fibrils (reviewed by Chiti and Dobson, 2006; Fowler *et al.*, 2007; Greenwald and Riek, 2010; Otzen and Nielsen, 2008).

Protein	Species	Function
	Bacterial amyloids	
Curli, Tafi	*E. coli* and other *Enterobacteriaceae*	Adhesion, cell invasion, biofilm formation, binding to host proteins
FapC	*Pseudomonas*	Biofilm formation
Chaplins	*Streptomycetes coelicolor* and other *Streptomyces*	Assisting aerial hyphae formation via modulation of water surface tension
Microcin E492	*Klebselia pneumonia*	Bacteriocin, membrane pore-forming, amyloid form inactive
Harpins	*Xanthomonas* and other plant pathogenic bacteria	Destabilisation of plant membranes, induction of cell death

Protein	Species	Function
MTP	*Mycobacterium tuberculosis*	Adherence factor
	Eukaryotic amyloids	
Pmel17	Mammals	Scaffolding and Sequestration of toxic intermediates during melanin synthesis
Peptide hormones	Mammals	Sorting, storage and release of hormones
Chorion proteins	Insects and fish	Structure-bearing and protection of oocytes and developing embryo
Spidroins	*Nephila clavipes* (spider)	Structure-bearing, formation of spider silk
Hydrophobins	Most fungi	Modulation of surface attachment and aerial hyphae formation
	Prion proteins	
ECTO-NOX	Mammals	[NADP(H)] oxidation and protein disulfide-thiol interchange
CPEB	*Aplysia* (sea slug)	Memory storage, leading to long term synaptic changes
HET-s	*Podospora anserina*	Regulation of heterokaryon formation
Sup35	*Saccharomyces cerevisiae*	Regulation of translation termination, prion form inactive
Ure2p	*Saccharomyces cerevisiae*	Regulation of nitrogen catabolism, prion form inactive
Rnq1p	*Saccharomyces cerevisiae*	Enhancing the inducibility of other prions
Swi1p	*Saccharomyces cerevisiae*	Chromatin remodelling
Mot3	*Saccharomyces cerevisiae*	Regulation of transcription of cell wall remodelling genes, prion form inactive

Based on their plastic properties, functional amyloids are involved in a large variety of processes such as biofilm formation, protection, adhesion and invasion, as well as the development of aerial structures, regulation of melanin synthesis and hormone release, molecular memories and information transfer by protein-based genetic elements.

Functional amyloids are biologically optimised and their formation seems to be highly regulated (reviewed by Greenwald and Riek, 2010). Sophisticated mechanisms allow the directed assembly into uniform fibrils, lacking cytotoxic oligomeric or protofibrillar intermediates (**chapter 1.1.3**). Currently known, these mechanisms of aggregation control comprise: chaperone assisted protein remodelling, enzymatic modifications, fibrillisation in specialised enclosed vesicles, nucleator dependent assembly and “gatekeeper” residues.

In yeast a chaperone network controls propagation of fibril formation and their degradation. The chaperones enable prion functionality and prevent amyloid/prion associated toxicity (Allen *et al.*, 2005; Gokhale *et al.*, 2005; Uptain and Lindquist, 2002). Similarly, a complex network of proteins is involved in curli biogenesis comprising six

different proteins (Barnhart and Chapman, 2006). The polymerisation into fibrils is nucleator dependent (Hammar *et al.*, 1996) and controlled by "gatekeeper" residues (Wang *et al.*, 2010). Due to "gatekeeper" residues distinct segments of the protein are less aggregation prone (Rousseau *et al.*, 2006). Thus, these residues avoid uncontrolled self-assembly, mislocalisation of the fibrils and cytotoxicity (Wang *et al.*, 2010). In the case of Pmel17 a proteolytic activation and stepwise modification is necessary (Berson *et al.*, 2001; Hoashi *et al.*, 2005). In addition, the fibrillisation of Pmel17 is localised to special vesicles in the melanosomes and depends on the pH (Berson *et al.*, 2001).

First and foremost, functional amyloids are widespread in bacteria. Recent studies show that amyloid adhesins are abundant in natural biofilms, in which at least 5 - 40 % of the prokaryotes produced amyloids, depending on their habitat (Larsen *et al.*, 2007). The curli fibrils of *E. coli* (Vidal *et al.*, 1998), Tafi fibrils of *Salmonella* (Austin *et al.*, 1998), TasA fibrils of *Bacillus subtilis* (Romero *et al.*, 2010) and FapC fibrils of *Pseudomonas* (Dueholm *et al.*, 2010) have clearly been identified as such bacterial adhesins. For example curli of *E. coli* form a loose fibrillar network on the surface of the bacteria and are a major component of the bacterial extracellular matrix. They assist biofilm formation, adhesion to host cells and inert surfaces, host cell invasion and the binding to host proteins (reviewed by Barnhart and Chapman, 2006). Functional amyloids and prions are also abundant in fungi and yeast (reviewed by Gebbink *et al.*, 2005; Wickner *et al.*, 2007). Analysis of the heterokaryon incompatibility in *Podospora anserina* identified the first functional prion – the HET-s prion (Coustou *et al.*, 1997). The systems of HET-s and curli are presented in detail in **chapters 1.2** (HET-s) and **1.3** (curli).

1.2 HET-s – A Prion Controls Heterokaryon Incompatibility

HET-s is one of the protein-based genetic elements in fungi (Wickner et al., 2008b). After solving the genetic puzzle of the filamentous ascomycetes *Podospora anserina* (Rizet, 1952), Begueret and co-workers reported HET-s as the first prion that displays a normal cellular function (Coustou *et al.*, 1997). HET-s is involved in a controlled cell death called heterokaryon incompatibility, which is suggested to prevent unlimited spread of fungal viruses between fusing colonies.

1.2.1 The Phenomenon of Heterokaryon Incompatibility

Heterokaryon incompatibility is an ubiquitous phenomenon in filamentous fungi (Glass and Dementhon, 2006). Fungal hyphae fuse when two colonies grow together. The result is a fusion cell (heterokaryon) with nuclei and cytoplasm of both parental colonies. This heterokaryon formation is genetically controlled and can be aborted by a programmed cell death. Only strains with certain identical loci can undergo hyphal fusion (Glass *et al.*, 2000; Saupe, 2000).

In *Podospora anserina* this mechanism (**Figure 1-5**) is based on the *het* locus with the genes *het-s* and *het-S* (Rizet, 1952; Turcq *et al.*, 1991). The encoded proteins HET-s and HET-S are isoforms. They are made of 289 amino acids and only differ in 13 residues. HET-s is a prion protein that forms amyloid fibrils resulting in a conversion of the non-prion phenotype [HET-s*] to the prion phenotype [HET-s]

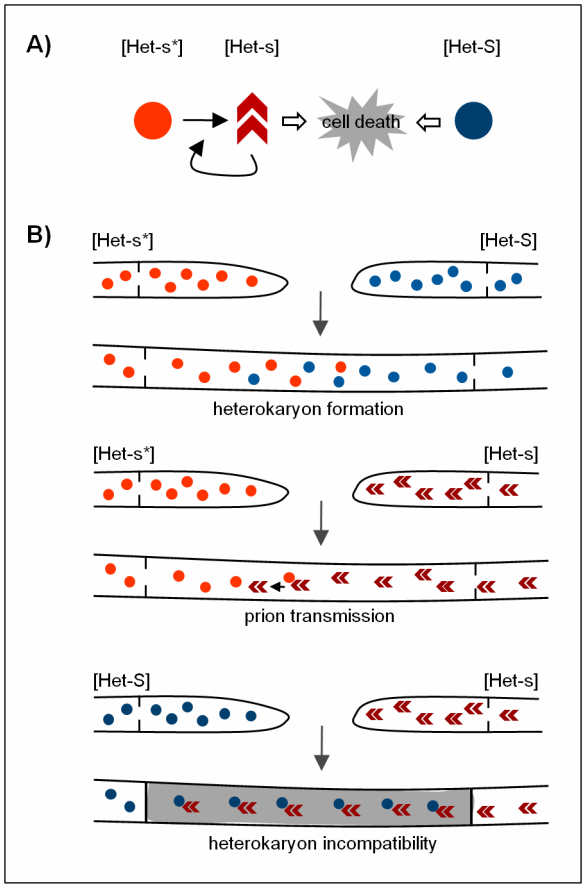

Figure 1-5: The Het-s/HET-S incompatibility. (A) The HET-s protein can exist in a soluble inactive form in [HET-s*] strains and in the prion form in [HET-s] strains. Interaction of the HET-s prion with HET-S leads to cell death. **(B)** Schematic representation of possible cell fusion events during interaction of strains differing in genotype and phenotype (adapted from Saupe, 2011; for explanation see text)

(Coustou *et al.*, 1997; Maddelein *et al.*, 2002). In the [HET-S] phenotype the HET-S protein is soluble and not amyloidogenic.

[HET-s*] strains are compatible with [HET-S] strains and cell fusion enables heterokaryon formation. The fusion of [HET-s*] and [HET-s] strains also leads to heterokaryon formation. Subsequently to the fusion, the [HET-s*] phenotype converts into [HET-s] via prion transmission and self-propagation. The fusion of a [HET-s] strain with a [HET-S] strain does not result in heterokaryon formation and finally the fusion cell dies. These corresponding phenotypes are incompatible. The dead cells act as a barrier between the growing colonies. In other words, the incompatibility is triggered by the interaction of HET-S with the prion form of HET-s (Greenwald *et al.*, 2010; Saupe, 2011).

1.2.2 Structure-Function Relationship of HET-s and HET-S

Podospora anserina heterokaryon incompatibility is driven by HET-s and HET-S – two proteins that share high sequence identity but differ especially in their ability to form amyloid fibrils (Greenwald *et al.*, 2010). Both proteins differ only in 13 amino acids (**Figure 1-6 A**), but only HET-s is able to form self-perpetuating amyloid fibrils in the prion state (Coustou *et al.*, 1997; Dos Reis *et al.*, 2002; Maddelein *et al.*, 2002). HET-s is a two-domain protein that consists of an N-terminal globular domain (residues ~ 1 – 227) and a C-terminal prion forming domain (pfd, residues 218 – 289) illustrated in **Figure 1-6 B** (Balguerie *et al.*, 2003). The N-terminal domain dictates the incompatibility type HET-s or HET-S and two amino acid differences at positions 23 and 33 are sufficient to cause incompatibility (Deleu *et al.*, 1993). In the prion state the pfd assembles into well ordered aggregates (Balguerie *et al.*, 2003; Dos Reis *et al.*, 2002).

Structural data are available for both domains and enable insights in the structure-function relationship of HET-s and HET-S. The crystal structures of the globular domain of both proteins revealed a HeLo domain with an α-helical fold of 8 – 9 helices with a short two-stranded β-sheet presented in **Figure 1-6 C** (Greenwald *et al.*, 2010; Idnurm and Howlett, 2003). This HeLo domain was identified to form dimers. HET-s and HET-S differ in the thermodynamic stability of their HeLo domain and HET-s constructs are thermodynamically more stable (Greenwald *et al.*, 2010).

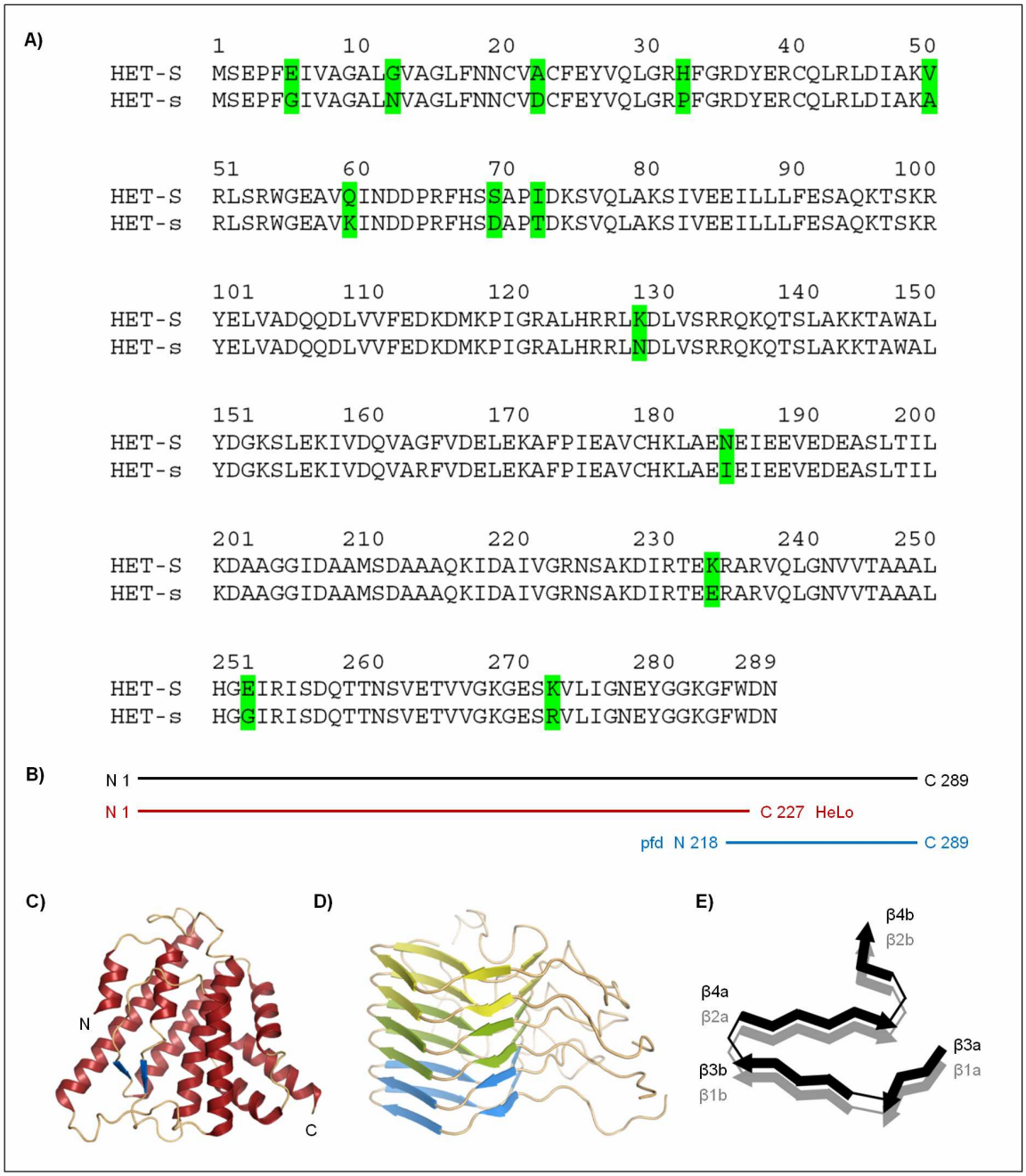

Figure 1-6: Domain organisation and structure of HET-s/HET-S. (A) Sequence alignment of HET-s and HET-S. Differing amino acids are highlighted in green. **(B)** Domain organisation of HET-s/HET-S. **(C)** Ribbon diagram of the HET-S HeLo domain (pdb entry 2VWO) (Greenwald *et al.*, 2010). **(D)** Side view of three stacked HET-s$_{218-289}$ molecules (pdb entry 2KJ3) (Van Melckebeke *et al.*, 2010; Wasmer *et al.*, 2008). Strands shown in the same colour correspond to one molecule. **(E)** Schematic representation of the two layers of the amyloid core formed by HET-s$_{218-289}$.

The first structural characterisation of the HET-s pfd via hydrogen/deuterium exchange NMR and solid state NMR spectroscopy (see **chapters 1.4.4** and **1.4.5**) identified four protected segments assigned to β-strands (Ritter *et al.*, 2005; Siemer *et al.*, 2005; Siemer *et al.*, 2006b). Subsequently, a structure with atomistic resolution was solved by solid state NMR and revealed further insights into the pfd structure shown in **Figure 1-6 D** (Van

Melckebeke *et al.*, 2010; Wasmer *et al.*, 2008). Thus, HET-$s_{218-289}$ forms a left-handed β-solenoid structure with a hydrophobic triangular core and two windings per monomer. Each winding is made of four short β-strands (β1a/β1b and β2a/β2b, β3a/β3b and β4a/β4b) shown in **Figure 1-6 E**. In each case two β-strands (a/b) are connected by sharp β-arcs and correspond to one previously assigned β-strand. While β2b and β4b point outwards, the first three strands of every turn contribute to the triangular core. The two β-strand segments are connected by a 15 residue flexible loop (Siemer *et al.*, 2006a). The topology of the core region is characterised by hydrophobic residues facing inside and polar residues facing outside. Furthermore, the structure is stabilised by three salt bridges and two asparagine ladders (Van Melckebeke *et al.*, 2010; Wasmer *et al.*, 2008). Former proline mutagenesis studies already indicated that this *in vitro* structure is highly similar to the prion structure formed *in vivo* (Ritter *et al.*, 2005).

Continuative studies allowed the proposal of a model (**Figure 1-7**) for the HET-s/HET-S incompatibility and induction of toxicity. This model is based on the local unfolding of the globular HeLo domain during the assembly of the pfd (Wasmer et al., 2009b), the tendency to form dimers of this globular domain as well as the proven interactions of HET-S with HET-s fibrils *in vivo* (Greenwald *et al.*, 2010). HET-S is suggested to interact with the HET-s fibril, which forces the HET-S pfd into the β-solenoid fold. Thereby the globular HeLo domain unfolds, followed by the formation of an aggregation prone dimer of the HeLo domain. These interactions are believed to block further fibril growth and to activate the toxicity of HET-S (Greenwald *et al.*, 2010; Saupe, 2011).

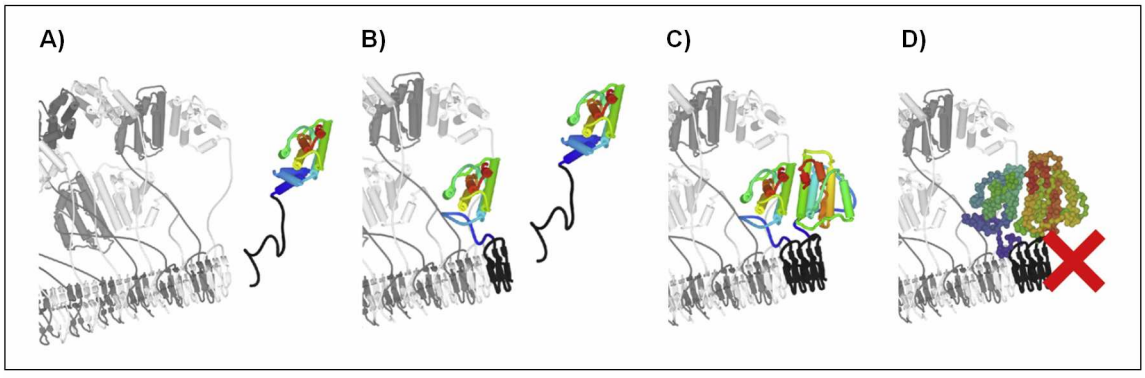

Figure 1-7: Proposed model of HET-s/HET-S incompatibility. The HET-s fibril is shown alternating in light and dark grey. Soluble HET-S is shown coloured (rainbow – HeLo domain, black – random coil pfd). **(A)** A soluble HET-S monomer reaches the HET-s fibril. **(B)** Via interaction of HET-S with the fibril edge the terminal α-helices of the HeLo domain unfold to allow the prion fold. **(C)** Via binding a second HET-S molecule the HeLo domains form a dimer. **(D)** Oligomerisation of the HeLo domain enables un/refolding and thus aggregation of the HeLo domain blocking further growth of the fibril (original figure from Greenwald *et al.*, 2010).

1.2.3 A HET-s Homologue Capable for Species Barrier Breaching

The analysis of the available fungal genomes (National Center for Biotechnology Information and Broad Fungal Genome Initiative) revealed several *Fusarium* species that feature HET-s homologous proteins with a putative prion forming domain (Wasmer *et al.*, 2010). In comparison to *Pa*HET-s, the homologue of *Fusarium graminearum Fg*HET-s exhibits highest sequence similarity. Both proteins show 50 % identity and 58 % homology overall illustrated in **Figure 1-8**. The protein segments, which correspond to the pfd, are identical at 37 % and homologous at 46 % of the sequence.

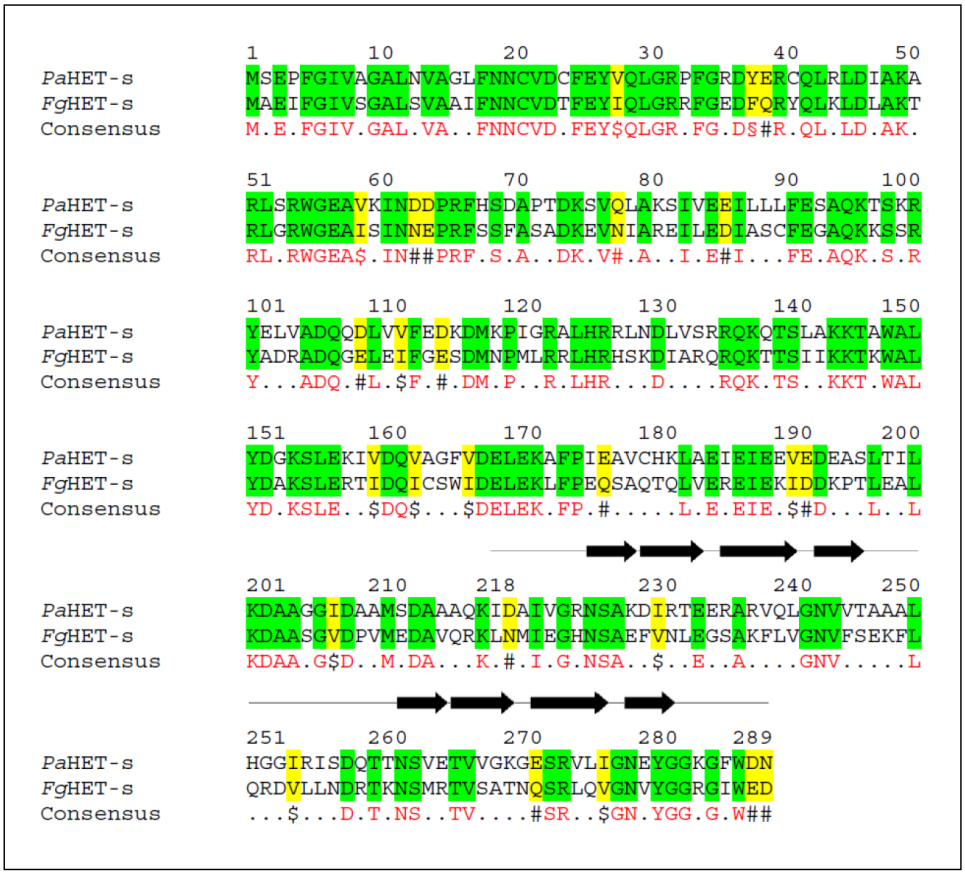

Figure 1-8: Sequence alignment of HET-s from *P. anserina* and *F. graminearum*. Primary structure of *Pa*HET-s is compared to *Fg*HET-s. Conserved residues are highlighted in green, homologous residues in yellow. In the consensus sequence conserved residues are shown in red, $ indicates anyone of I or V, % anyone of F or Y, # anyone of D, E, N or Q. For the prion forming domain the β-strands of the β-solenoid structure of *Pa*HET-$s_{218-289}$ Wasmer *et al.*, 2008 are shown on the top of the alignment. Analysis was done with the program MultAlin (Corpet, 1988).

Saupe and co-workers identified this homologue to be capable for cross-seeding in spite of extensive sequence divergence (Wasmer *et al.*, 2010). *In vitro Fg*HET-$s_{218-289}$ fibrils nucleate the amyloid fibril formation of *Pa*HET-$s_{218-289}$ and *vice versa* indicating species barrier breaching. These results are quite astonishing. So far prions were obtained to

transmit only rarely or not at all between different species, a phenomenon called species barrier (Prusiner, 1998; Wickner *et al.*, 2009). In prion transmission preformed amyloid fibrils act as structural template (Chien *et al.*, 2004; Cohen and Prusiner, 1998). The soluble protein adopts the same conformation as the template fibrils. Transmission and cross-seeding presuppose protein sequences and structures that are highly similar (Prusiner, 1998). An amino acid sequence dependent host range controls the prion transmission. But in the case of the cross-seeding between *Pa*HET-$s_{218-289}$ and *Fg*HET-$s_{218-289}$ the sequence identity (37 %) is comparatively very low. Certainly, amyloids have been shown to form different polymorphic forms and much different structures already due to single point mutations (Tycko *et al.*, 2009). Only globular, soluble proteins were assumed to adopt the same structure corresponding to a certain homology level (Chothia and Lesk, 1986; Ginalski, 2006). Thus, the question about the background of the observed cross-seeding effect between the two homologues arose. Obviously, the structure elucidation of the *Fg*HET-s prion for comparison with *Pa*HET-s is required to answer this question.

1.3 Curli Fibrils and Their Major Subunit CsgA

Curli amyloid fibrils are the main component of a complex extracellular matrix formed by many enterobacteria like *E. coli* and *Salmonella* (**Figure 1-9**). Initially curli fibrils were found in *E. coli* strains that caused bovine mastitis in the late 1980s (Olsén *et al.*, 1989). Recent studies identified curli as key player in adhesion to abiotic and biotic surfaces, host invasion, binding to host proteins and biofilm formation (reviewed by Barnhart and Chapman, 2006).

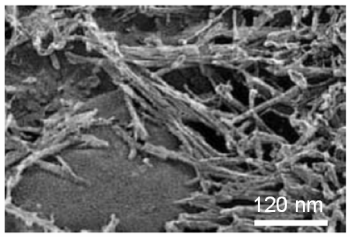

Figure 1-9: Curli on the surface of *E. coli* cells. Electron micrograph of curliated bacteria. (Chapman *et al.*, 2002).

In *E. coli* curli fibril formation six different proteins are involved. These proteins are encoded by the two operons *csgBA* and *csgDEFG* (Hammar *et al.*, 1995). The homologous fibrils in *Salmonella* are called Tafi (thin aggregative fimbrae) and the operons are *agfBA* and *agfDEFG* (Collinson *et al.*, 1996; Romling *et al.*, 1998a). The encoded proteins are highly conserved among the two species (Romling *et al.*, 1998a). The following section will describe the fibrils of both species and will be referred to as curli.

1.3.1 Curli Biogenesis Is Highly Controlled

Curli fibrils are built up of two highly homologous proteins – the major curli subunit CsgA and the nucleator protein CsgB. Fibril formation requires four additional proteins CsgC, CsgE, CsgF and CsgG, which interact in a sophisticated mechanism (**Figure 1-10**).

The operons *csgBA* and *csgDEFG* encode the corresponding curli specific genes. Curli gene expression is stringently regulated and responsive to many environmental cues (Gerstel and Romling, 2003). CsgD is the transcription factor and promotes the transcription of *csgBA* (Hammar *et al.*, 1995). Thus, the regulation of *csgD* expression also influences curli gene expression. Regularly, curli gene expression is enhanced by a growth temperature below 30 °C (Arnqvist *et al.*, 1992; Olsén *et al.*, 1989). However, high expression levels of curli specific genes were obtained in clinical *E. coli* strains at 37 °C (Bian *et al.*, 2000), and mutations in the *csgD* promoter region can lead to curli gene expression without any temperature dependence (Romling *et al.*, 1998b; Uhlich *et al.*, 2001). Furthermore, growth media without salt (Romling *et al.*, 1998b), limitation of nutrients like nitrogen, phosphate and iron (Gerstel and Romling, 2001; Romling *et al.*,

1998b) and microaerophilic conditions (Gerstel and Romling, 2001; Romling *et al.*, 1998a) are able to promote curli gene expression.

The curli specific gene products CsgA, CsgB, CsgE, CsgF and CsgG carry an N-terminal Sec sequence that enables the proteins to cross the inner membrane via a Sec system. CsgG mediates the secretion of CsgA and CsgB out of the cell (Loferer *et al.*, 1997; Robinson *et al.*, 2006) This outer membrane lipoprotein forms oligomeric ring-shaped complexes, similar to other outer membrane channel-forming proteins (Brok *et al.*, 1999; Thanassi and Hultgren, 2000; Thanassi *et al.*, 1998). CsgE and CsgF can interact with the membrane pore and are suggested to act as chaperones during curli biogenesis (Robinson *et al.*, 2006). While CsgE is believed to avoid premature CsgA assembly in the periplasm (Nenninger *et al.*, 2011), CsgF is also secreted to the extracellular space, where it facilitates the cell-association and protease resistance of CsgB (Nenninger *et al.*, 2009).

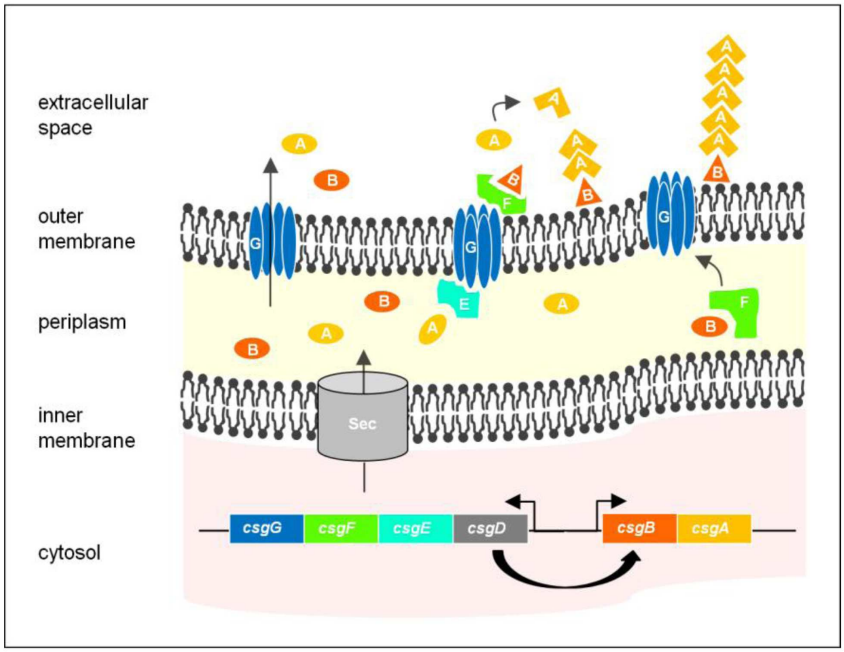

Figure 1-10: Curli biogenesis. The two curli gene operons are shown at the bottom. CsgD enhances the transcription of *csgBA*. All other gene products carry a sec signal sequence to enter the periplasm via the Sec pathway. CsgG is an outer membrane pore mediating the secretion of CsgA and CsgB. Outside the cell CsgB nucleates CsgA into well ordered fibrils. CsgE and CsgF interact with CsgG and facilitate curli assembly in a chaperone like manner (adapted from Barnhart and Chapman, 2006; Nenninger *et al.*, 2009).

After secretion of CsgA and CsgB, CsgB nucleates the polymerisation of soluble CsgA monomers into long unbranched fibrils resulting in a loose network of curli fibrils (Chapman *et al.*, 2002; Hammar *et al.*, 1996). Furthermore, the three internal repeats of the fivefold repetitive core of CsgA display nonconsensus residues (Wang *et al.*, 2010). These "gatekeeper" residues are aspartates and glycines, which reduce the aggregation grade of the distinct CsgA segments. Thus, these residues avoid uncontrolled self-assembly,

mislocalisation of the curli fibrils and cytotoxicity (Rousseau *et al.*, 2006; Wang *et al.*, 2010). Recent studies identified the operon *csgBA* in *Salmonella* to encode a third protein – CsgC (Gibson *et al.*, 2007). Currently, the role of CsgC in curli biogenesis is unclear, but loss of CsgC leads to aberrant fibril formation.

1.3.2 The Manifold Roles of Curli

Curliated bacteria benefit from the surrounding fibrillar network in many different ways. Next to cellulose and other polysaccharides, curli are the main component of the extracellular matrix formed by many enterobacteria (Zogaj *et al.*, 2003). This matrix protects the embedded cells and is an important requirement to form biofilms (Hall-Stoodley *et al.*, 2004; Van Houdt and Michiels, 2005). Biofilms are aggregates of microorganisms, in which the single cells benefit from an increased resistance to environmental stress. In addition, biofilms provide mechanical stability and metabolisable substrates. Since curli also allow the bacteria to adhere to Teflon and stainless steel, biofilm formation is often a problem in food industry and hospital settings (Austin *et al.*, 1998).

Furthermore, curli can bind to the extracellular matrix proteins fibronectin (Collinson *et al.*, 1993; Olsén *et al.*, 1989) and laminin (Olsén *et al.*, 1993), as well as to MHC class I, which presents antigens on the cell surface (Olsén *et al.*, 1998). Thus, curli support the attachment and invasion of host cells (reviewed by Barnhart and Chapman, 2006). In addition, these surface fibrils interact with other intracellular host proteins enhancing bacterial spreading throughout the host (Bian *et al.*, 2000). Curli can attach to plasminogen and the tissue type plasminogen activator. This attachment enables the activation of plasmin leading to the degradation of soft tissue. Thus, the bacteria can access to deeper tissue layers (Sjobring *et al.*, 1994). Interactions with the contact-phase proteins H-kininogen, fibrinogen and factor XII retard clotting and also enhance the dissemination of the curliated bacteria (Ben Nasr *et al.*, 1996; Herwald *et al.*, 1998; Olsén *et al.*, 2002). As pathogen-associated molecular pattern (PAMP), curli are recognised by Toll-like receptors inducing the innate immune system via proinflammatory cytokines (Froy, 2005; Kai-Larsen *et al.*, 2010; Tükel *et al.*, 2010). In summary, curli play an important role in a number of mechanisms of pathogens. Thus, detailed insights in the structure-function relationship of curli would enhance the understanding of bacterial pathogenesis.

1.3.3 The Major Curli Subunit CsgA

Curli fibrils consist of the elongator protein CsgA, which is the major curli subunit, and the nucleator protein CsgB. Both proteins share 30 % sequence identity and 40 % homology (Hammar *et al.*, 1995). *In vivo* CsgA assembles into well ordered amyloid fibrils only in the presence of CsgB (Hammar *et al.*, 1996; Hammer *et al.*, 2007). In contrast to the *in vivo* assembly, *in vitro* formation of curli-like fibrils does not require CsgB (Chapman *et al.*, 2002; Wang *et al.*, 2007).

The amino acid sequence of CsgA comprises a signal peptide for protein export (residues 1 – 20), 22 flexible amino acids (residues 21 – 43) at the mature N-terminus and an amyloid core (residues 44 – 151), which is made of five imperfect repeats (**Figure 1-11**). Hence, mature CsgA comprises both, a protease-susceptible and a protease-resistant domain (Collinson *et al.*, 1999). The amyloid core is characterised by intramolecular highly conserved residues. Every repeated unit shows the consensus sequence of $Sx_5QxGxGNxAx_3Qx_{1\text{-}5}$ and consists of 19 – 23 amino acids (Collinson *et al.*, 1999; Hammar *et al.*, 1996). The conserved glutamine and asparagine residues are suggested to form a hydrogen bonding network that stabilises the fibrils of CsgA (Collinson *et al.*, 1999). Chemically synthesised peptides of the units 1, 3 and 5 are also amyloidogenic (Wang *et al.*, 2007). Recombinantly expressed and *in vitro* polymerised mature CsgA shows all the characteristics of amyloid fibrils. CsgA fibrils are resistant to proteases and SDS and show a β-sheet rich fold (Chapman *et al.*, 2002; Shewmaker *et al.*, 2009; Wang *et al.*, 2007). In the presence of polymerised CsgA, ThT shows an increased fluorescence signal, allowing the visualisation of the fibrillisation kinetics (Wang *et al.*, 2007).

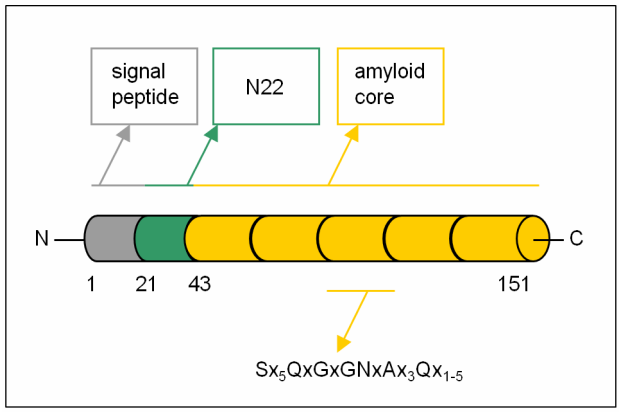

Figure 1-11: Domain structure of CsgA. Mature CsgA consists of a flexible N-terminus and an amyloid core made of five imperfect repeats with a characteristic consensus sequence (x is any amino acid).

A first study on the structure of CsgA with solid state NMR spectroscopy indicated that the amyloid core is not made of a parallel in-register β-sheet structure (Shewmaker *et al.*, 2009). The authors concluded that a β-helix like structure would be likely, but the obtained data were not sufficient to identify this structure.

1.4 Techniques to Study Amyloid Structures

The three-dimensional structure of every protein dictates its function. Thus, protein structure elucidation is indispensable in order to understand the corresponding functions. Amyloid fibrils form filamentous structures, which are inherently difficult to study. Conventional methods for protein structure-function analysis and thermodynamic characterisation are usually not feasible for amyloids. Amyloids do not form crystals at all, which are required for X-ray crystallography. Multi-subunit proteins in the megadalton range such as amyloids are too large for solution NMR spectroscopy. Recent research enhanced the known techniques for structure analysis to overcome the challenging properties of amyloids (reviewed by Chiti and Dobson, 2006; Toyama and Weissman, 2011), as for example hydrogen/deuterium (H/D) exchange NMR.

1.4.1 Low Resolution Techniques

For a long time only imaging techniques like TEM and AFM, as well as X-ray fibril diffraction analysis revealed insights into amyloid structures as illustrated in **chapter 1.1.2** (Harper *et al.*, 1997; Sunde *et al.*, 1997). Fourier transform infrared (FT-IR) and circular dichroism (CD) spectra are helpful to verify a β-sheet rich fold by specific absorption bands, as well as the detection of the folding kinetics (Berthomieu and Hienerwadel, 2009; Ranjbar and Gill, 2009). Moreover, FT-IR spectra identify amyloid specific β-sheets by a characteristic wave number of the amide I band at 1625 – 1610 cm^{-1}, compared to soluble proteins with a wave number ≥ 1630 cm^{-1} (Zandomeneghi *et al.*, 2004). Extrinsic dyes like ThT and CR (**chapter 1.1.2**) are sensitive tools to probe folding intermediates, surface hydrophobicity and aggregation or fibrillisation of proteins (Hawe *et al.*, 2008). As most fibrils bind to these specific dyes, they can be used to demonstrate the tertiary and quaternary structure of the amyloid fold as well.

Initially, these methods were complemented by techniques like proteolytic digestion (Kheterpal *et al.*, 2001) and mutagenesis studies with the β-breaker proline (Chiba *et al.*, 2003; Williams *et al.*, 2004; Wood *et al.*, 1995) to identify stable regions. An alternative mutagenesis approach has been established by using the introduction of single cysteine residues. The thiol group of a cysteine residue can be linked to paramagnetic spin labels for electron paramagnetic resonance (EPR) measurements (Chen *et al.*, 2007; Der-Sarkissian *et al.*, 2003; Torok *et al.*, 2002) and extrinsic fluorescence dyes (Krishnan and Lindquist,

2005; Ohhashi *et al.*, 2010; Ritter *et al.*, 2005). The success of the mutagenesis analysis depends on the correct folding of the overall structure of the modified proteins compared to the wild type form. In general, corresponding experiments provide information about the presence or absence of regular secondary structure elements at a defined residue, the topology of β-sheets, as well as inter- and intramolecular contacts within the fibril.

1.4.2 High Resolution Techniques

Further developments in biophysical methods facilitate a closer insight into amyloid structures by cryo electron microscopy, X-ray crystallography, as well as NMR spectroscopy techniques as H/D exchange NMR or solid state NMR.

Cryo electron microscopy (cryo EM) is an advanced form of electron microscopy at cryogenic temperatures. Electrons that transmit through the sample are detected (Henderson, 2004). Cryo EM enables the analysis of amyloids under physiological conditions without staining or fixation, which is necessary in conventional EM. The observed electron densities indicate the helical twist of the protofilaments and the cross-β structure of amyloids (Jimenez *et al.*, 1999; Jimenez *et al.*, 2002; Meinhardt *et al.*, 2009; Sachse *et al.*, 2008; Schmidt *et al.*, 2009; Zhang *et al.*, 2009).

Eisenberg and co-workers were able to use X-ray crystallography to analyse microcrystals of peptide fragments forming amyloid-like structures. In several studies they revealed atomistic details congruent to existing structural models of amyloids (Nelson *et al.*, 2005; Sawaya *et al.*, 2007; Wiltzius *et al.*, 2009; Wiltzius *et al.*, 2008). Thus, the β-sheet stacking interface “dry steric zipper” was postulated by Eisenberg as a basic principle of the cross-β structure illustrated in **chapter 1.1.2** (Nelson *et al.*, 2005; Sawaya *et al.*, 2007). Since the analysed peptides are only up to seven residues, the wealth of data needs to be confirmed by techniques probing the full-length protein.

1.4.3 Theoretical Background of NMR Spectroscopy

Another approach to reveal atomistic resolution of the protein structure is based on nuclear magnetic resonance (NMR) spectroscopy. NMR spectroscopy is widely used for the structure elucidation of small molecules, but it can also be applied for structure determination and structure-function relationship analysis of proteins under physiological

conditions (Wüthrich, 1990). In addition to conventional structure determination of proteins, NMR enables for example the investigation of folding intermediates (Hofmann *et al.*, 2008; Schulenburg *et al.*, 2009), chemical properties of functional groups (Gardiennet *et al.*, 2005; Shammas *et al.*, 2007), intermolecular interactions (Debnath *et al.*, 2011; Drechsler *et al.*, 2011), hydrogen bonding (Assadi-Porter *et al.*, 2003; Liwang and Bax, 1997) as well as structural investigation of amyloids and membrane proteins (Hoshino *et al.*, 2002; Linser *et al.*, 2011; Tycko, 2011).

All these NMR applications are based on the phenomenon of nuclear magnetic resonance of certain nuclei. Most nuclei have an intrinsic angular momentum *P*. This angular momentum depends on the nuclear quantum number *I* shown in **Equation 1-1**.

Equation 1-1: $$P = \sqrt{I(I+1)} \cdot \frac{h}{2\pi}$$ h = Planck constant

The nuclei have an intrinsic angular momentum *P*, whenever I ≠ 0. Placed in an external magnetic field B_0, the intrinsic angular momentum *P* is orientated along B_0. The energy level of the nucleus is split in 2*I*+1 energy states shown in **Figure 1-12** (Zeeman effect, Zeeman, 1897).

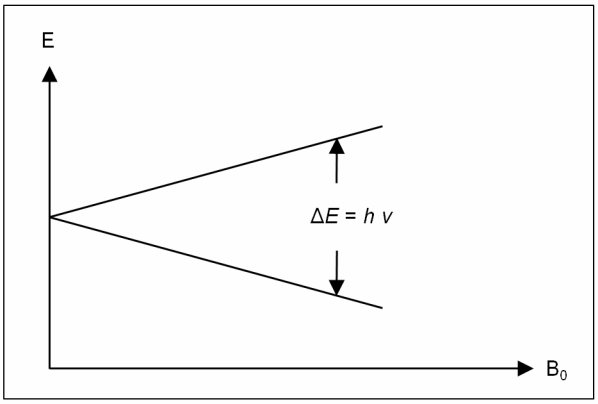

Figure 1-12: Zeeman effect of a nucleus with *I* = ½. The energy level of a nucleus with *I* = ½ splits into two energy states (2*I*+1).

The application of a radio frequency pulse B_1, whose energy is equal to ΔE of the split energy states of the nucleus, leads to the saturation of the higher energy level. The relaxation of the higher energy level can be detected. Depending on the chemical environment, the magnetic field B_0 on a nucleus is shielded by the B_0-induced magnetic field of the neighbouring nuclei and their electrons. The result is the local magnetic field B_{eff}. Higher shielding effects result in a decreased B_{eff}. For a constant B_0 the resonance frequency decreases. Different nuclei are distinguished by their resonance frequencies, which are influenced by the chemical environment. The different resonance frequencies are

given in the chemical shift δ in ppm. The resonance frequency of the nuclei ν_{sample} is divided by the resonance frequency ν_0 referred to a reference signal ν_{ref} for example tetramethylsilane (TMS) as illustrated in **Equation 1-2**.

Equation 1-2:
$$\delta = \frac{\nu_{sample} - \nu_{ref}}{\nu_0} \cdot 10^6$$

Important nuclei in NMR spectroscopy are nuclei with I=1/2 ^{1}H, ^{13}C, ^{15}N, ^{19}F, ^{31}P. The natural abundance of ^{13}C and ^{15}N nuclei is relative low compared to ^{1}H isotopes. For this reason, proteins chosen for NMR experiments have to be labelled with these isotopes.

The NMR technique has been further developed to multi-dimensional (2D, 3D) NMR experiments to overcome the overlap of signals in 1D NMR spectra. These NMR experiments are necessary to solve the complex structure of proteins. Multi-dimensional data sets are generated from series of one-dimensional experiments and cross-signals (cross-peaks) are analysed. Accordingly [^{15}N, ^{1}H] correlation spectra for example reveal only cross-peaks of NH-groups, because the corresponding experiments detect exclusively the scalar couplings between H- and N-nuclei.

1.4.4 The H/D Exchange NMR Experiment

H/D exchange [^{15}N, ^{1}H] NMR spectroscopy in combination with a sequence specific assignment of resonances is a powerful tool to identify structured regions in proteins at single residue resolution (Englander, 2000; Li and Woodward, 1999). This experiment is based on the different hydrogen exchange rates of backbone amides that depend on their conformation. Deuterium is not detectable in the [^{15}N, ^{1}H] NMR spectrum.

Hydrogen exchange occurs every time between the protons of a protein and the surrounding solvent (Englander *et al.*, 1997; Hvidt and Linderstrøm-Lang, 1954). Influenced by the chemical environment, the exchange is retarded (Bai *et al.*, 1993). Backbone amides, which are involved in hydrogen bonds, are less accessible to the solvent. The hydrogen exchange reaction is slowed down. Backbone amides, which are not involved in hydrogen bonding, exchange faster. Recording NMR spectra provides information about the protonation of backbone amides and their exchange behaviour. Since secondary structure elements are formed by hydrogen bonding, this exchange behaviour allows to the distinction between ordered and disordered regions of a protein.

Amyloids are not feasible for conventional NMR experiments as they form megadalton complexes. However, dissolving them in aprotic dimethyl sulfoxide (DMSO) enables the analysis of unfolded monomers of amyloid proteins after H/D exchange experiments as reported by Goto and co-workers (Hirota-Nakaoka *et al.*, 2003; Hoshino *et al.*, 2002). The H/D exchange NMR experiment of amyloids (**Figure 1-13**) includes three distinct steps (Hoshino *et al.*, 2002):

A) Incubation of the amyloid fibrils in D_2O buffer

B) Monomerisation of the fibrils with DMSO

C) NMR analysis.

At first the amyloid fibrils are incubated in D_2O buffer to enable H/D exchange between the protein and the solvent. The reaction is quenched at appropriate time intervals by removing the solvent. Secondly, the fibrils are monomerised in DMSO. Furthermore, the aprotic DMSO conserves the hydrogen/deuterium pattern present in the fibril at a specific incubation time. Finally, [^{15}N, ^{1}H] correlation spectra are recorded and analysed regarding the presence and absence of resonances. In a spectrum of fibrils that were incubated in water each resonance corresponds to one backbone amide of the protein. Hydrogen/deuterium exchange results in a spectrum with significantly reduced or absent resonances. The backbone amides corresponding to the absent resonances have undergone a fast exchange and deuterium is not detectable in such NMR spectroscopy.

Furthermore the H/D exchange NMR technique has been improved in different ways (Luhrs *et al.*, 2005; Ritter *et al.*, 2005). One improvement is the measurement of series of [^{15}N, ^{1}H] correlation spectra in order to correct the resonance intensities for exchange in DMSO. Remaining traces of D_2O are able to mask the resonance intensities due to further exchange in DMSO. For this reason, a rapid mixing with DMSO and preadjusted NMR settings (together less than 1 min) are of great importance. The acquisition of one spectrum was reduced to less than 5 min. Another aspect is the analysis of NMR spectra after different incubation intervals in D_2O. These data enable the accurate examination of exchange rates of each assigned backbone amide. Finally, the exchange rates provide additional information about the homogeneity and regularity within the fibrils (Luhrs *et al.*, 2005; Ritter *et al.*, 2005).

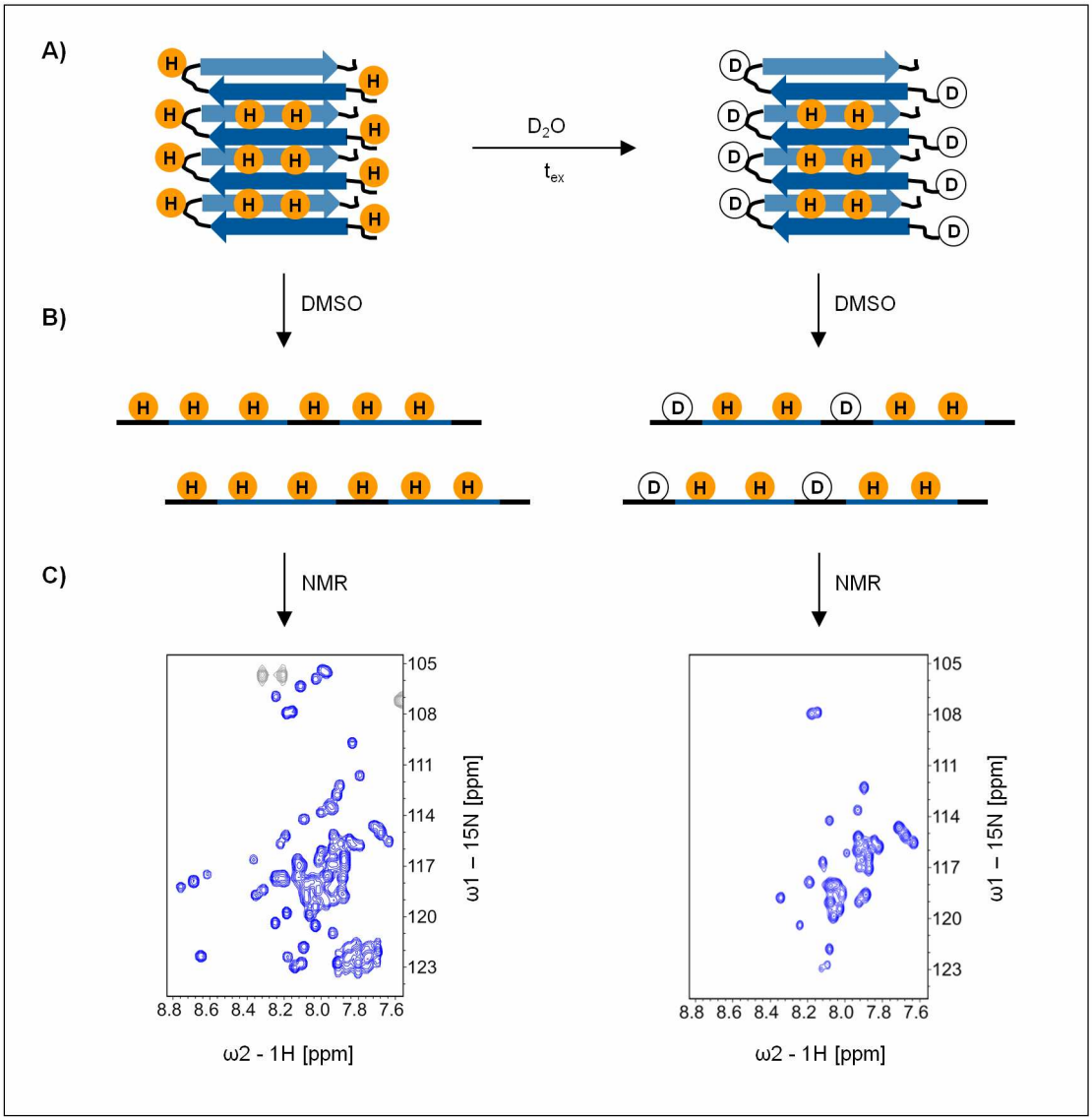

Figure 1-13: Schematic description of the H/D exchange NMR experiment. (A) H/D exchange within the fibril. Reaction is started by solvent exchange to D_2O. Hydrogens involved in hydrogen-bonds forming secondary structure are protected, contrary hydrogens of flexible backbone amides are easily substituted by solvents deuterons. **(B)** Monomerisation and unfolding of the amyloid fibrils using DMSO. Original positions of β-strands are highlighted in the same colour as in (A). Dissolving the fibrils in aprotic DMSO conserves the proton/deuteron pattern. **(C)** Detection of H/D exchange via NMR recording [^{15}N, ^{1}H] correlation spectra. Obtained NMR spectra differ in their peak intensities.

In summary, H/D exchange NMR spectroscopy is a unique tool to study stability and thermodynamic behaviour of proteins and their variants. This approach was already successful applied to a number of systems to map the amyloid cores of HET-s (Ritter *et al.*, 2005), Sup35 (Toyama *et al.*, 2007), Aβ (Luhrs *et al.*, 2005), β_2-microglobulin (Hoshino *et al.*, 2002), α-synuclein (Vilar *et al.*, 2008) and SH3 domain fibrils (Carulla *et al.*, 2005). Alternatively, H/D exchange experiments have been evaluated by mass spectrometry HD-MS (Kheterpal *et al.*, 2006). The higher mass of deuterium also enables the detection

of the exchange by mass spectrometry. H/D exchange and quenching the reaction are followed by the digestion of the protein into smaller peptides. Subsequently the digested peptides are separated and analysed by mass spectrometry (Kheterpal *et al.*, 2006). One advantage of HD-MS experiments is the requirement of less protein material. One major drawback is the missing ability to digest the protein into small peptides suitable for precise analysis and the artifactual exchange during this digestion. As many amyloids are only soluble in hexafluoroisopropanol, formic acid or DMSO, it is difficult to find a solvent suitable for monomerisation and adjacent digestion by a protease. For the prion forming domain of HET-s from *Podospora anserina* H/D exchange experiments were analysed and compared for both techniques illustrated in **Figure 1-14** (Nazabal and Schmitter, 2006; Ritter *et al.*, 2005). The comparison indicated obvious differences in the resolution of the techniques.

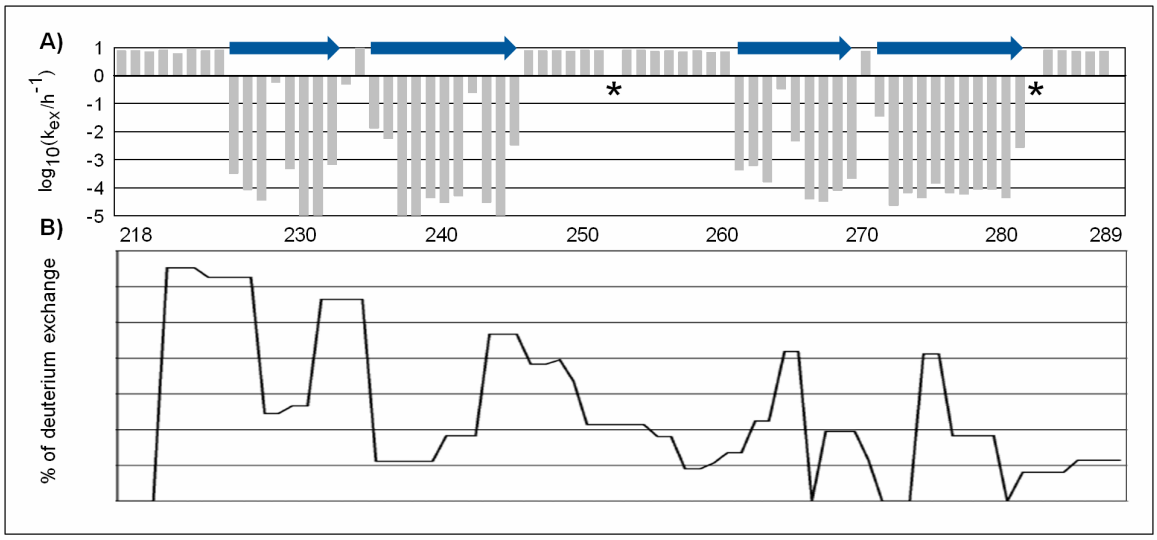

Figure 1-14: H/D exchange analysis of HET-s$_{218-289}$. (A) NMR analysis identified four protected regions assigned to β-strands (Ritter *et al.*, 2005). Bars indicate hydrogen exchange rates k_{ex} (h^{-1}), blue arrows indicate β-strands. Slow exchange rates correspond to backbone amides, which are involved in hydrogen bonding. Black asterisks mark residues, which could not be analysed. **(B)** MS analysis determined the percentage (%) of deuterium incorporation along the HET-s$_{218-289}$ sequence (Nazabal and Schmitter, 2006). Less deuterium incorporation indicates protected backbone amides.

1.4.5 Solid State NMR Spectroscopy

In contrast to solution NMR spectroscopy, solid state NMR (ssNMR) experiments allow direct structural elucidation of solid material. The NMR samples in ssNMR experiments are rotated at high frequency at 54.74 ° to the magnetic field, called magic angle spinning (Andrew, 1981; Hennel and Klinowski, 2005), in order to increase the resolution and

decrease anisotropy and dipole dipole couplings effects (Heise, 2008). This technique provides information about polypeptide mobility, interresidue distances and residue-specific positions of secondary structure elements (Heise, 2008; Tycko, 2000, 2006, 2011). Furthermore, solid state NMR experiments enable protein structure elucidation at atomic resolution. In the early 2000s Oschkinat and co-workers could solve the first complete protein structure by ssNMR (Castellani *et al.*, 2002). Tycko and co-workers were the first, who applied ssNMR to amyloids with their work on the amyloid β peptide $A\beta_{1-40}$ (Antzutkin *et al.*, 2000; Balbach *et al.*, 2002; Petkova *et al.*, 2002; Petkova *et al.*, 2006). In addition to $A\beta_{1-40}$, this technique was successfully applied to determine the β-sheet architecture for example of Sup35 (Shewmaker *et al.*, 2006), α-synuclein (Heise *et al.*, 2005), Rnq1p (Wickner et al., 2008a) and Ure2p fibrils (Baxa *et al.*, 2007). Unfortunately uniformly labelled proteins often lead to crowded spectra with overlapping peaks. This fact leads to imprecise assignments and absent distance information. Nevertheless, with HET-s the first amyloid structure could recently be solved, discussed in **chapter 1.2.2** (Wasmer *et al.*, 2008).

2 Aims and Scope

The HET-s prion of *Podospora anserina* and the curli system of *E. coli* are two of the most studied systems that form functional amyloids (reviewed by Barnhart and Chapman, 2006; Saupe, 2011). They constitute a group of unique systems that allows the analysis of how proteins fibrillise without forming toxic intermediates resulting in a physiologically valuable fold. In order to understand this phenomenon with regard to distinct functions, structural information is essential.

Remarkable results acquired Saupe and co-workers in cross-seeding experiments of two HET-s homologues: In spite of the low sequence identity, *Fg*HET-s is capable to nucleate *Pa*HET-s fibrillisation and *vice versa*. In order to analyse the origin of the unexpected HET-s cross-seeding phenomenon, one aim of this work was to investigate the structural relationship of the homologous prion domains. Therefore, *Fg*HET-$s_{218-289}$ should be subjected to H/D exchange NMR spectroscopy to investigate the structural regularity, homogeneity and stability of the fibrils in comparison to *Pa*HET-s. Since the structure of the *Pa*HET-s prion domain was already well characterized, H/D exchange NMR should be probed for its capability to compare the topology of the homologues without a complete solid state NMR analysis for both proteins. For the meaningful determination of residue-specific H/D exchange rates, *Fg*HET-$s_{218-289}$ had to be recombinantly produced and specifically fibrillised, the sequence-specific backbone-resonance assignment was to be established.

The curli fibrils of *Enterobacteria* were shown to facilitate manifold functions in bacterial pathogens. Since CsgA is the major curli subunit, its structure-function relationship is of high interest. As there was no residue-specific information about structural elements of CsgA, another aim of this thesis was the structural characterisation of CsgA from *E. coli* following two strategies: the analysis of specific fibrils as well as the analysis of amyloid-like folding intermediates. In order to investigate the topology, as well as homogeneity and stability of *Ec*CsgA fibrils, H/D exchange NMR experiments and fluorescence labelling were intended. For this purpose, new protein variants, protein production, specific fibrillisation and fluorescence labelling as well as the backbone-resonance assignment had to be arranged. *In vivo* the fibrillisation of CsgA is nucleator dependent. Therefore, heteronucleated *Ec*CsgA fibrils should be probed by H/D exchange NMR to evaluate the influence of the nucleator *Ec*CsgB on the structure, homogeneity and stability of *Ec*CsgA.

In order to analyse amyloid like folding intermediates of *Ec*CsgA, stable soluble analogues of *Ec*CsgA oligomers are required. Therefore, appropriate expression constructs, as well as the recombinant production of corresponding proteins and purification protocols had to be established. Probably stable soluble protein constructs should be analysed with regard to their secondary/tertiary structure and capability for structural analysis.

3 Material and Methods

3.1 Standard Materials

3.1.1 Chemicals, Enzymes, Antibodies and Kits

If not stated otherwise, all chemicals that were used in this work were of *"pro analysis"* grade. Chemicals were purchased from the companies Amersham Biosciences, Fluka, Merck, Millipore, Omnilab, Qiagen, Roche, Roth, Sigma and Stratagene.
Enzymes (**Table 3-1**), antibodies (**Table 3-2** and **Table 3-3**) and kits (**Table 3-4**) were used as described by the corresponding company.

Table 3-1: Enzymes used in this work.

Enzyme	Source	Specification
Hen white-egg lysozyme	Fluka	Glycoside hydrolase
DNAse	Merck	Endodeoxyribonuclease
Proteinase K	EuroClone	Serine peptidase
Shrimp Alkaline Phosphatase	Roche	Phosphate hydrolase
Platinum *Pfx* Polymerase	Invitrogen	DNA polymerase
Pfu Turbo Polymerase	Stratagene	DNA Polymerase
T4 DNA Ligase	New England Biolabs	DNA Ligase
BamHI, DpnI, HindIII, NotI, SacI, SapI, TatI, XbaI	New England Biolabs, Fermentas	Restriction endonucleases

Table 3-2: Primary antibodies used in this work.

Name	AB species	Antigen	Dilution	Source
anti-His6	Mouse	Hexa-His	1:1500	Novagen

Table 3-3: Secondary antibodies used in this work.

Name	AB species	A/R species	Conjugate	Dilution	Source
Anti Mouse IgG (H+L) AP Conjugate	Goat	Mouse	Alkaline phosphatase	1:7500	Promega
Anti Mouse IgG (H+L) HRP Conjugate	Goat	Mouse	Horse radish peroxidase	1:2000	Dianova

Table 3-4: Kits used in this work.

Kit	Source	Usage
QIAprep® Spin Miniprep Kit	Qiagen	Plasmid purification

Kit	Source	Usage
QIAquick® PCR Purification Kit	Qiagen	Purification of PCR products, removal of nucleotides
QIAquick® Gel Extraction Kit	Qiagen	DNA extraction from gel slices
QuikChange® Site-Directed Mutagenesis Kit	Qiagen	Selective modification of plasmid DNA by mutation
BugBuster Kit® Protein Extraction Reagent	Novagen	Protein extraction after expression test
BCIP/NBT Color Development Substrate	Promega	Detection of AP conjugated antibodies
Lumi-Light Western Blotting Substrate	Roche	Detection of HRP conjugated antibodies by ECL

3.1.2 Molecular Weight Standards

Following molecular weight standards were used (**Table 3-5**).

Table 3-5: Molecular weight standards used in this work.

Name	Source	Type
Smart Ladder	Eurogentec	Agarose gel DNA marker
Unstained Protein Molecular weight marker	Fermentas	SDS-PAGE protein marker
PageRuler™ Unstained Protein ladder	Fermentas	SDS-PAGE protein marker
PageRuler™ Prestained Protein ladder	Fermentas	SDS-PAGE protein marker

3.1.3 Bacterial Strains

Following bacterial strains were used (**Table 3-6**).

Table 3-6: Bacterial strains used in this work.

E. coli strain	Genotype	Source
BL21 (DE3)	*F⁻ ompT gal dcm lon hsdS$_B$(r_B^- m_B^-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	Stratagene
DL39	*F⁻ Lam⁻acpC13 fnr-25rph-1 ilve12 tyrB507*	LeMaster and Richards, 1988
DH5α	*F⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(r_K^- m_K^+), λ–*	Invitrogen
T7 Express	*fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10--Tet^S)2 [dcm] R(zgb-210::Tn10--Tet^S) endA1 Δ(mcrC-mrr)114::IS10*	NEB
TUNER (DE3)	*F⁻ lacY ompT gal dcm lon hsdS$_B$(r_B^- m_B^-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	Novagene
XL1-Blue	*rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *lac*	Stratagene

3.1.4 Plasmids

Following recombinant plasmids were used in this work (**Table 3-7**). Plasmid derivatives of pET-11d, pET-15b, pETRO, pRSET-A mT and pBar100 carried an ampicillin resistance gene. pET-24a and pETM-10 carried a kanamycin resistance gene. In addition to the ampicillin resistance, pBar100 plasmids carried a chloramphenicol resistance gene too.

Table 3-7: Recombinant plasmids used.

Plasmid	Features	Source
pRSET-A mT (Zahn *et al.*, 1996)	N-terminal hexa histidine tag, engineered thrombin cleavage site	T. Lührs; HZI, Braunschweig
pBar100 (Schmoldt *et al.*, 2005)	pMT416 derivative encoding barnase peptide fusion constructs	M. Daneschdar/H. Kolmar; TU Darmstadt
pET-11d $CsgA_{21-151}$	pET-11d derivative encoding *E. coli* $CsgA_{21-151}$	A. Eberth; HZI, Braunschweig
pET-11d $CsgA_{21-151}$ V23C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ E44C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ L45C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ N46C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ S55C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ L57C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ A58C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ L59C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ D80C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ V81C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ S88C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ S89C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ S100C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ A125C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ V126C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ S134C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ V135C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ N136C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ T147C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ A148C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{21-151}$ H149C	Site directed mutagenesis of pET-11d $CsgA_{21-151}$	this work
pET-11d $CsgA_{40-151}$	pET-11d derivative encoding *E. coli* $CsgA_{40-151}$	this work
pRSET-A mT CsgA CM	pRSET-A mT derivative encoding fusion protein R5c-CsgA-R1c	this work

Plasmid	Features	Source
pET-15b CsgA CM	pET-15b derivative encoding fusion protein R5c-CsgA-R1c	this work
pET-15b CsgA CM V25K/L161K	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM N26K/N162K	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM T37K/L173K	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM A38K/A174K	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM V25K/N26K/L161K/N162K	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM T37K/A38K/L173K/A174K	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM V25P/L161P	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM T28P/Y164P	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM N35P/S171P	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM A38P/A174P	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM V25P/N35P/L161P/S171P	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM T28P/A38P/Y164P/A174P	Site directed mutagenesis of pET-15b CsgA CM	this work
pETRO CsgA CM	pETRO derivative encoding fusion protein R5c-CsgA-R1c	this work[1)]
pRSET-A mT CsgA CD	pRSET-A mT derivative encoding fusion protein R5c-$(CsgA)_2$-R1c	this work
pET-15b CsgA CD	pET-15b derivative encoding fusion protein R5c-$(CsgA)_2$-R1c	this work
pET-15b CsgA CM V25K/L274K	Site directed mutagenesis of pET-15b CsgA CD	this work
pET-15b CsgA CM N26K/N275K	Site directed mutagenesis of pET-15b CsgA CD	this work
pET-15b CsgA CM T37K/L286K	Site directed mutagenesis of pET-15b CsgA CD	this work
pET-15b CsgA CM A38K/A287K	Site directed mutagenesis of pET-15b CsgA CD	this work
pET-15b CsgA CM V25K/N26K/L274K/N275K	Site directed mutagenesis of pET-15b CsgA CD	this work
pET-15b CsgA CM T37K/A38K/L286K/A287K	Site directed mutagenesis of pET-15b CsgA CD	this work
pET-15b CsgA CM V25P/L274P	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM T28P/Y277P	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM N35P/S284P	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM A38P/A287P	Site directed mutagenesis of pET-15b CsgA CM	this work
pET-15b CsgA CM V25P/N35P/L274P/S284P	Site directed mutagenesis of pET-15b CsgA CM	this work

Plasmid	Features	Source
pET-15b CsgA CM T28P/A38P/Y277P/A287P	Site directed mutagenesis of pET-15b CsgA CM	this work
pETRO CsgA CD	pETRO derivative encoding fusion protein R5c-$(CsgA)_2$-R1c	this work[1)]
pET-11d Sumo-CsgA-Sumo	pET-11d derivative encoding fusion protein Sumo-CsgA-Sumo	this work
pET-11d Bar-$(CsgA)_2$-Bar	pET-11d derivative encoding fusion protein Bar-$(CsgA)_2$-Bar	this work
pET-11d Sumo-$(CsgA)_2$-Sumo	pET-11d derivative encoding fusion protein Sumo-$(CsgA)_2$-Sumo	this work
pET-24a *Fg*HET-$s_{218-289}$	pET-24a derivative encoding *F. graminearum* HET-$s_{218-289}$	C. Wasmer; ETH Zürich

[1)] The original pETRO vector was established by T. Lührs/S. Loss (HZI, Braunschweig) and kindly provided by them.
Sequences of encoded proteins (excepting variants due to site directed mutagenesis) are given in the Appendix.

3.1.5 Oligonucleotides

Oligonucleotides (**Table 3-8**) were used as primers for site-directed mutagenesis and PCR. They were purchased from MWG.

Table 3-8: Oligonucleotides used. Recognition sites for restriction enzymes are underlined and point mutations/codon exchanges are marked in red. For QuikChange site-directed mutagenesis only forward primers are listed.

Name	Sequence (5´ → 3´)	Restriction site/ Mutation/Tag
	Cloning and sub-cloning	
*csg*A·ΔN39 fwd	TCATGCCATGGGCCCAAATTCTGAGC	*Nco*I
*csg*A·ΔN39 rev	GGCTTTGTTAGCAGCCGGATCCTTAG	*Bam*HI
*csg*A·fwd	GTAATGGATCCAATAGCGGCCCAAAT	*Bam*HI
*csg*A·rev	CTCAGAAGCTTTTAGTACTGATGAGCG	*Hind*III
*csg*A·SacI	GGCCCAAATTCTGAGCTCAACATTTACCAGTACGGT	*Sac*I
*csg*A·R5c	GTTCCGCGTGGATCCAATAGCGGCCCAAATAGCAGCGTTAACGTTACCCAGGTGGGTTTCGGCAATAATGCCACAGCCCACCAATATTCTGGTGATAACTCTGAGCTCAACATT	*Bam*HI, *Sac*I
*csg*A·R1c	CCAACAGTACTCTGGCGACAATAGCGAATTGAATATCTATCAATATGGCGGTGGCAATAGCGCCCTGGCACTTCAGACCTAAAAGCTTGATCCG	*Tat*I, *Hind*III
*csg*A·ΔTatI	GACCGCTCATCAGTATTCCGGCGACAACAGC	Δ*Tat*I
*csg*A·R1R1c fwd	TTAGAGTACTCCGGTGATAACTCT	*Tat*I
*csg*A·R1R1c rev	GGATCAAGCTTTTAGGTCTGAAGT	*Hind*III
pRSET·ΔTatI	TCAGAATGACTTGGTTGAGTATTCACCAGTCACAGAAAAAGC	Δ*Tat*I
bar·N fwd	TCATGCCATGGGCGCACAGGTTATCAACAC	*Nco*I

Name	Sequence (5´ → 3´)	Restriction site/ Mutation/Tag
bar·N rev	AATGTTGAGCTCAGAATTTGGGCCGCTTCTGATTTTTGTAAAGGTC	*Sac*I
*csg*A·fus fwd	TCATGCCATGGGCAGCGGCCCAAATTCTGAG	*Nco*I, *Sac*I
*csg*A·fus rev	CGCGGATCCTTAGTGATGGTGATGGTGATGGCTCATGGAAGAGCTGTACTGATGAGCGGTCG	*Bam*HI, *Sap*I
bar·C fwd	CTTGCTCTTCCATGAGCGCACAGGTTATCAACAC	*Sap*I
bar·C rev	CGCGGATCCTTAGTGATGGTGATGGTGATGTCTGATTTTTTGTAAAGGTC	*Bam*HI, H6
sumo·N fwd	TCATGCCATGGGTCATCACCATCATCATCACGGGTCCCTACAGGAC	*Nco*I, H6
sumo·N rev	AATGTTGAGCTCAGAATTTGGGCCGCTCCTGCCAGGCATGC	*Sac*I
sumo·C fwd	GAAGCTCTTCCATGAGCGGGTCCCTACAGGAC	*Sap*I
sumo·C rev	CGCGGATCCTTACCTGCCAGGCATGCCACCAAT	*Bam*HI
*csg*Ad·fwd	TTAGAGTACTCTGGTGACAACTCTGAGCTGAACATTTACCAGTACGGTG	*Tat*I, Δ*Sac*I
*csg*Ad·rev	GGATCAAGCTTGCTCATGGAAGAGCTATACTGATGAGCGGTCG	*Hind*III, *Sap*I
pET11d·Δ SapI	AGGAAGCGGAAGGGCGCCTGATGCG	Δ*Sap*I
Mutagenesis to create Cys variants of *Ec*CsgA		
*csg*A·V23C	ATATACCATGGGTGTTTGTCCTCAGTACGGCGGCG	V23C
*csg*A·E44C	GTAATAATAGCGGCCCAAATTCTTGCCTGAACATTTACCAGTACGGTGG	E44C
*csg*A·L45C	AGCGGCCCAAATTCTGAGTGCAACATTTACCAGTACGGTGG	L45C
*csg*A·N46C	CCCAAATTCTGAGCTGTGCATTTACCAGTACGGTGGC	N46C
*csg*A·S55C	GGTGGCGGTAACTGTGCACTTGCTCTGC	S55C
*csg*A·L57C	GGCGGTAACTCTGCATGTGCTCTGCAAACTGATGC	L57C
*csg*A·A58C	GGTGGCGGTAACTCTGCACTTTGTCTGCAAACTGATGCCCGTAAC	A58C
*csg*A·L59C	GGTAACTCTGCACTTGCTTGCCAAACTGATGCCCGTAACTC	L59C
*csg*A·D80C	CGGCGGTAATGGTGCATGTGTTGGTCAGGGCTCA	D80C
*csg*A·V81C	CGGTAATGGTGCAGATTGTGGTCAGGGCTCAGATG	V81C
*csg*A·S88C	CAGGGCTCAGATGACTGCTCAATCGATCTGACC	S88C
*csg*A·S89C	GGCTCAGATGACAGCTGTATCGATCTGACCCAACGT	S89C
*csg*A·S100C	CGTGGCTTCGGTAACTGCGCTACTCTTGATCAG	S100C
*csg*A·A125C	GTGGTGGCAACGGTGCTTGCGTTGACCAGACTGCATC	A125C
*csg*A·V126C	TGGCAACGGTGCTGCATGTGACCAGACTGCATCTA	V126C
*csg*A·S134C	ACTGCATCTAACTCCTGCGTCAACGTGACTCAGG	S134C
*csg*A·V135C	CTGCATCTAACTCCTCCTGCAACGTGACTCAGGTTGG	V135C

Name	Sequence (5´ → 3´)	Restriction site/ Mutation/Tag
*csg*A·N136C	CTGCATCTAACTCCTCCGTCTGCGTGACTCAGGTTG GCTTTGG	N136C
*csg*A·T147C	TGGCTTTGGTAACAACGCGTGCGCTCATCAGTACCA TCACC	T147C
*csg*A·A148C	GCTTTGGTAACAACGCGACCTGTCATCAGTACCATC ACCATC	A148C
*csg*A·H149C	GGTAACAACGCGACCGCTTGTCAGTACCATCACCAT CACC	H149C
Mutagenesis to create Lys and Pro variants of CsgA CM and CsgA CD		
V25K	GCGGCCCAAATAGCAGCAAGAACGTTACCCAGGT GGG	V25K
N26K	CCCAAATAGCAGCGTTAAAGTTACCCAGGTGGGTT	N26K
T37K	GTTTCGGCAATAATGCCAAAGCCCACCAATATTCT GG	T37K
A38K	GGGTTTCGGCAATAATGCCACAAAGCACCAATATT CTGGTGATAACT	A38K
L161K L274K	TATTCTGGCGACAATAGCGAAAAGAATATCTATCA ATATGGCGGTG	L161K L274K
N162K N275K	GGCGACAATAGCGAATTGAAAATCTATCAATATGG CGGT	N162K N275K
L173K L286K	CGGTGGCAATAGCGCCAAGGCACTTCAGACCTAAA	L173K L286K
A174K A287K	GGTGGCAATAGCGCCCTGAAACTTCAGACCTAAAA GCTT	A174K A287K
V25K/N26K	TAGCGGCCCAAATAGCAGCAAGAAAGTTACCCAG GTGGGTTTCG	V25K/N26K
T37K/A38K	CAGGTGGGTTTCGGCAATAATGCCAAAAAGCACCA ATATTCTGGTGATAACTCT	T37K/A38K
L161K/N162K L274K/N275K	TATTCTGGCGACAATAGCGAAAAGAAAATCTATCA ATATGGCGGTGGC	L161K/N162K L274K/N275K
L173K/A174K L266K/A287K	TCAATATGGCGGTGGCAATAGCGCCAAGAAACTTC AGACCTAAAAGCTT	L173K/A174K L266K/A287K
V25P	CGGCCCAAATAGCAGCCCGAACGTTACCCAGGTGG	V25P
T28P	AGCAGCGTTAACGTTCCGCAGGTGGGTTTCGGC	T28P
N35P	GGTGGGTTTCGGCAATCCGGCCACAGCCCACCAAT	N35P
A38P	GTTTCGGCAATAATGCCACACCGCACCAATATTCT GGTGATAA	A38P
L161P, L274P	TATTCTGGCGACAATAGCGAACCGAATATCTATCA ATATGGCGGT	L161P, L274P
Y164P, Y277P	GACAATAGCGAATTGAATATCCCGCAATATGGCGG TGGCAATAGC	Y164P,Y277P
S171P, S284P	AATATGGCGGTGGCAATCCGGCCCTGGCACTTCAG A	S171P,S284P
A174P, A287P	GGCAATAGCGCCCTGCCGCTTCAGACCTAAAAG	A174P, A287P

3.2 Media and Buffer

The media and buffer, which were used in this study, are summarised in **Table 3-9** and **Table 3-10**. All media were sterilised by autoclaving (121 °C, 2 bar, 20 min, Top7000PST, Sauter) and heat-sensitive additives were sterile filtered (pore width 0.2 µM). Antibiotics were added after media cooling below 50 °C. Depending on plasmid and bacterial strain the following antibiotics were used: ampicillin (100 µg/mL), chloramphenicol (34 µg/mL) and kanamycin (30 µg/mL).

Table 3-9: Media used for bacterial culture.

Medium	Composition
Lysogeny Broth (LB)	1 % (w/v) tryptone, 0.7 % (w/v) NaCl, 0.5 % (w/v) yeast extract, adjusted to pH 7.5 with NaOH
LB-Agar	LB, 1.5 % (w/v) agar-agar
Terrific Broth (TB)	1.2 % (w/v) tryptone, 2.4 % (w/v) yeast extract, 0.4 % (v/v) glycerol, 0.1 x TB salts
TB salts	0.17 M KH_2PO_4, 0.72 M K_2HPO_4
Minimal Medium (M9)	0.2 x M9 salts, 19 mM NH_4Cl, 0.4 % (w/v) glucose, 1 mM $MgSO_4$, 0.0002 x trace metal mix, 0.01 x MEM vitamin solution (SigmaAldrich)
M9 salts	240 mM Na_2HPO_4, 110 mM KH_2PO_4, 1 % (w/v) NaCl
Trace metal mix	50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnSO_4$, 10 mM $ZnSO_4$, 2 mM $CoCl_2$, 2 mM $CuCl_2$, 2 mM $NiCl_2$, 2 mM Na_2MoO_4, 2 mM H_3BO_3
Muchmore Medium (MMM9) (Muchmore *et al.*, 1989)	0.05 % (w/v) Ala, 0.04 % (w/v) Arg, 0.04 % (w/v) Asp, 0.005 % (w/v) cystine, 0.04 % (w/v) Gln, 0.065 % (w/v) Glu, 0.055 % (w/v) Gly, 0.01 % (w/v) His, 0.023 % (w/v) Ile, 0.023 % (w/v) Leu, 0.042 % (w/v) Lys · HCl, 0.025 % (w/v) Met, 0.013 % (w/v) Phe, 0.01 % (w/v) Pro, 0.21 % (w/v) Ser, 0.023 % (w/v) Thr, 0.023 % (w/v) Val 0.05 % (w/v) adenine, 0.065 % (w/v) guanosine, 0.02 % (w/v) thymine, 0.05 % (w/v) uracil, 0.02 % (w/v) cytosine, 18.3 mM NaAc, 0.15 % (w/v) succinic acid, 9.3 mM NH_4Cl, 21.3 mM NaOH, 77 mM KH_2PO_4, 0.05 x supplementary solution, adjusted to pH 7.5 with NaOH
Supplementary solution	40 % (w/v) glucose, 80 mM $MgSO_4$, 0.2 mM $FeCl_3$, 0.27 mM $CaCl_2$, 0.14 mM $ZnSO_4$, 0.24 mM $MnSO_4$, 0.1 % (w/v) Trp, 3.3 mM thiamine, 8.1 mM niacin, 80 µM biotin

Table 3-10: Solution and buffers used in this work.

Solution/Buffer	Composition
AP buffer	100 mM Tris/HCl pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$
Blocking solution	TBST, 5 % (w/v) skimmed milk powder
Coomassie staining solution	30 % (v/v) ethanol, 10 % (v/v) acetic acid, 0.25 % (w/v) Coomassie R-250
Coomassie destaining solution	40 % (v/v) ethanol, 10 % (v/v) acetic acid
Congo red stock solution	0.2 % (w/v) congo red, 1 % (w/v) NaCl, 80 % (v/v) ethanol, 0.001 % (w/v) NaOH
6 x DNA loading buffer	10 mM Tris/HCl pH 7.4, 25 mM EDTA, 30 % (v/v) glycerol,

Solution/Buffer	Composition
	0.4 % (w/v) OrangeG
IB wash buffer A (*Ec*CsgA)	50 mM Tris/HCl pH 8, 1.5 M NaCl, 1 % (w/v) N-Lauroylsarcosine
IB wash buffer B (*Fg*HET-s)	50 mM Tris/HCl pH 8, 1.5 M NaCl, 1 % (v/v) Triton X-100
4 x Lower buffer	1.5 M Tris/HCl pH 8.8
Lysis buffer I	50 mM Tris/HCl pH 8, 150 mM NaCl
Lysis buffer II	50 mM Tris/HCl pH 7.5, 150 mM NaCl, 5 % (v/v) glycerol, 0.1 % (v/v) Triton X-100, 0.01 % (w/v) SDS, 0.05 % (w/v) N-Lauroylsarcosine
Native buffer A (*Ec*CsgA)	50 mM KPi pH 7.2
Native buffer B (*Fg*HET-s)	50 mM Tris/HCl pH 8, 150 mM NaCl
Ponceau S staining solution	2 % (w/v) Ponceau S, 30 % (w/v) TCA, 30% (w/v) sulfosalicylic acid
Proteolysis buffer	125 mM Tris/HCl pH 7.5, 0.8 M urea
Y x SDS sample buffer	0.1 M Tris/HCl pH 6.8, 2 % (w/v) SDS, 0.2 % (w/v) Bromphenol blue, 20 % (v/v) glycerol, 5 mM DTT
SDS separating gel (15 %)	7.5 mL Rotiphorese® Gel 30 (Carl Roth), 3.8 mL 4 x lower buffer, 0.15 mL 10 % (w/v) SDS, 3.55 mL MilliQ, 0.02 mL TEMED, 0.03 mL 25 % (w/v) APS
SDS separating gel (12 %)	6 mL Rotiphorese® Gel 30 (Carl Roth), 3.8 mL 4 x lower buffer, 0.15 mL 10 % (w/v) SDS, 5 mL MilliQ, 0.02 mL TEMED, 0.05 mL 25 % (w/v) APS
SDS stacking gel (5 %)	0.8 mL Rotiphorese® Gel 30 (Carl Roth), 1.3 mL 4 x upper buffer, 3 mL MilliQ, 7.5 µL TEMED, 12.5 µL 25 % (w/v) APS
Solubilisation buffer A (*Ec*CsgA)	100 mM KPi pH 7.2, 150 mM NaCl, 8 M GdnHCl
Solubilisation buffer B (*Fg*HET-s)	50 mM Tris/HCl pH 8, 500 mM NaCl, 6 M GdnHCl, 5 mM imidazole
1 x TAE buffer	40 mM Tris, 20 mM sodium acetate, 1 mM EDTA, adjusted to pH 8.2 with acetic acid
1 x TBST	20 mM Tris/HCl pH 8, 150 mM NaCl, 0.05 % (v/v) Tween-20
Transfer buffer for Western Blotting	20 mM Tris pH 8, 192 mM glycine, 15 % (v/v) Methanol
4 x Upper buffer	0.5 M Tris/HCl pH 6.8, 0.4 % (w/v) SDS
Urea/SDS	8 M urea, 2 % (w/v) SDS

3.3 Microbiological and Molecular Biological Methods

Molecular biology methods, which were applied in this study, were adapted from standard procedures (Ausubel *et al.*, 2007; Coligan, 2003; Sambrook and Russell, 2000) and are not described in detail unless essential modifications were made.

3.3.1 General Conditions of Bacterial Culture

In order to cultivate bacteria for gene expression and plasmid amplification, cells were always freshly transformed with plasmid DNA by heat shock or electroporation. Subsequently, the transformed cells were incubated overnight on agar plates that contained an antibiotic, which corresponded to the used plasmid. For liquid cultures an appropriate volume of culture medium was inoculated with a single colony. These cultures were incubated in baffled flasks under shaking (120 – 180 rpm, depending on number of baffles and ventilation). All bacterial cultures were grown at 20 - 37 °C.
Plasmids were extracted from DH5α and XL1-Blue cells. BL21(DE3), DL39, T7 Express and TUNER(DE3) cells were subjected to bacterial culture for recombinant gene expression.

3.3.2 DNA Analytical Methods

Agarose gel electrophoresis was performed to analyse and purify DNA samples after PCR, modification by restriction enzymes and mutagenesis. The gels were prepared from 0.8 or 2 % (w/v) agarose in TAE buffer and supplemented with 5 ‰ (v/v) ethidium bromide. DNA samples were mixed with 6 x DNA loading buffer. The gels were run in TAE buffer with 5 V/cm for 60 min and documented under UV illumination at 254 nm. The yield and purity of isolated bacterial plasmid DNA, PCR products and modified DNA during cloning were checked spectrophotometrically by absorption at 260 and 280 nm. The DNA concentration was calculated from A_{260}, assuming that A_{260} = 1 is equivalent to 50 µg/mL DNA. With a ratio of A_{260}/A_{280} of 1.8 – 2 DNA was supposed to be pure.

3.3.3 Molecular Cloning

During this work certain protein variants were analysed and several cloning steps were necessary to establish the different constructs. The enzymes, bacterial strains, plasmids and oligonucleotides, which were employed to this purpose, are summarised in **Table 3-1**, **Table 3-6**, **Table 3-7** and **Table 3-8**. If not stated otherwise, all reactions were carried out as described in the protocols of the manufacturers' instructions.
In general, genes of interest were amplified by PCR using the Platinum *Pfx* Polymerase and the amplificates were purified by agarose gel electrophoresis and adjacent agarose gel extraction. Afterwards the obtained polynucleotides as well as the recipient plasmids were digested with the corresponding restriction enzymes. Linearised plasmids were dephosphorylated by Shrimp Alkaline Phosphatase to avoid self-ligation. In order to remove the modifying enzymes and remaining nucleotides, the PCR amplificates and the

linearised plasmids were applied to the PCR purification kit. In preparation of ligation, plasmid- and target-DNA were mixed in a molar ratio of 1:5 and heated to 90 °C for 5 min. This mixture was cooled down and incubated with T4 DNA Ligase at 16 °C overnight and finally 2 µL of the reaction volume were utilised for transformation.
In order to modify plasmid-DNA selectively, QuikChange® Site-Directed Mutagenesis was performed using the *Pfu* Turbo Polymerase.
All constructs were confirmed by sequencing in the department of Genome Analytics (HZI, Braunschweig) using a 3100-Avant genetic analyser (ABI Prism) or by the companies GATC (Konstanz) and MWG (Ebersberg).

3.3.4 Generation of Plasmid Constructs to Obtain Fusion Proteins

A schematic representation of gene constructs encoding CsgA fusion proteins is shown in **Figure 3-1**.
Gene constructs that encode CsgA fusion proteins with additional edge strands (CsgA CM and CsgA CD) were cloned into the pRSET-A mT vector (Zahn *et al.*, 1996) after removing the *Tat*I recognition site in the plasmid backbone (pRSET·ΔTatI). First *csg*A$_{38\text{-}151}$ was amplified using the primer pair *csg*A·N fwd, *csg*A·C rev and cloned into the given vector by *Bam*HI and *Hind*III. A silent mutation was introduced via *csg*A·SacI generating a *Sac*I recognition site. In order to fuse additional edge strands, the ss-oligonucleotide pairs of *csg*A·R1c and *csg*A·R5c were annealed by heating to 95 °C (1 °C/min, start 20 °C) and cloned into pRSET-A mT CsgA$_{38\text{-}151}$ using *Bam*HI, *Sac*I and *Tat*I and *Hind*III respectively. Thus, pRSET-A mT R5c-*csg*A$_{38\text{-}151}$-R1c was generated, named pRSET-A CsgA CM.
To generate a gene construct, which encodes CsgA CD, *csg*A$_{38\text{-}151}$-R1c was again amplified using the primer pair *csg*A·R1R1c fwd, *csg*A·R1R1c rev and pRSET-A mT R5c-*csg*A$_{38\text{-}151}$-R1c as template after removing the *Tat*I recognition site by a silent mutation (*csg*A·ΔTatI). The obtained PCR product was introduced in a pRSET-A CsgA CM vector with an intact *Tat*I recognition site in the *csg*A gene using *Tat*I and *Hind*III resulting in pRSET-A CsgA CD. Using *Xba*I and *Hind*III, the CsgA CM and CsgCD constructs were also cloned into the vectors pET-15b and pETRO (established and provided by T. Lührs/S.Loss; HZI, Braunschweig).
The oligonucleotides of *csg*A·R1c and *csg*A·R5c were designed in that way that their sequence differs from the original DNA sequence but is still encoding for the same amino acids. This fact allowed specific modifications of R1c and R5c by QuikChange® Mutagenesis.
Gene constructs that encode barnase- and sumo-fusion proteins of CsgA were established using the pET-11d vector as follows: *csg*A$_{38\text{-}151}$ was amplified using pRSET-A mT CsgA CM as template and the primer pair *csg*A·fus fwd, *csg*A·fus rev. The obtained PCR product was cloned into pET-11d via *Nco*I and *Bam*HI. For N- and C-terminal fusion of barnase the corresponding gene was amplified based on pBar100 via *bar*·N fwd, *bar*·N rev and *bar*·C fwd, *bar*·C rev respectively and introduced into the obtained pET-11d derivative using *Nco*I, *Sac*I and *Sap*I, *Bam*HI, respectively. Cloning of sumo constructs was done accordingly using 0912068_new_3Sumostar_for_Ch_pMA as template (Geneart construct, kindly provided by J. van den Heuvel, HZI, Braunschweig).

To generate fusion constructs, which include a duplicate of the five CsgA repeats, first a "dimer" had to be created. Therefore intermediates of establishing pRSET-A CsgA CD could be used. pRSET-A mT CsgA CM lacking *Tat*I in *csg*A was used to amplify a *csg*A construct flanked by *Tat*I and *Sap*I, *Hind*III recognition site respectively (*csg*Ad·fwd, *csg*Ad·rev). The obtained product was cloned into pRSET-A mT R5c-*csg*A$_{38\text{-}151}$-R1c by *Tat*I and *Hind*III resulting in a duplicate of the five CsgA repeats, which was accordingly introduced into pET-11d bar-csgA-bar and pET-11d sumo-csgA-sumo by *Sac*I and *Sap*I.

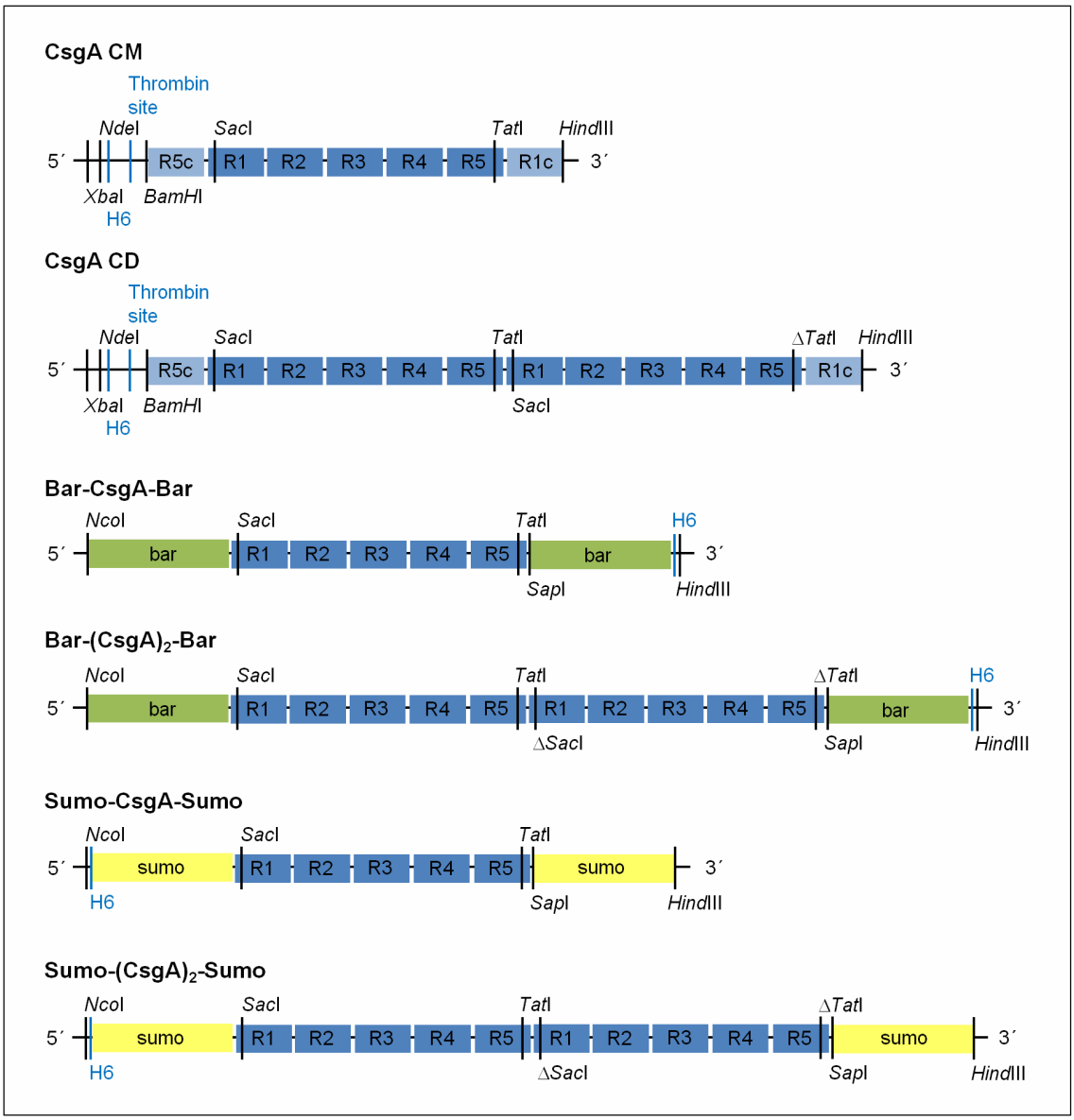

Figure 3-1: Cloning sites for gene constructs encoding CsgA fusion proteins. CsgA CM und CsgA CD present *Ec*CsgA constructs with additional edge strands (R1c and R5c shown in light blue), Bar-(CsgA)$_n$-Bar and Sumo-(CsgA)$_n$-Sumo present CsgA constructs fused to barnase (green) or sumo (yellow). CsgA repeating units (R1 – R5) are shown in blue, DNA restrictions sites, protease cleavage sites and affinity tags are highlighted. The coding sequences are shown in 5´ → 3´ direction.

3.4 Protein Production and Purification

3.4.1 Test Expression

Small scale test expressions were performed to establish the recombinant expression of novel gene constructs. Thus, 20 mL liquid medium were inoculated with a single colony and cultures were grown to an OD_{600} of 0.6 – 0.8 at 37 °C. Optimal IPTG concentration, expression time and expression temperature were verified by varying these parameters. At several time points after induction samples were taken by V (µL) = $300/OD_{600}$. These samples were further analysed for soluble and insoluble expressions products using the Bug-Buster® Kit followed by SDS PAGE. Previous to SDS PAGE, insoluble fractions were solubilised with 8 M urea or 98 % (v/v) formic acid. Formic acid was removed using a speed vac at 60 °C for 1 h and the pellets were dissolved in SDS sample buffer as described in **chapter 3.5.4**.

3.4.2 Recombinant Protein Synthesis

According to recombinant expression, bacteria were cultivated as described in **chapter 3.3.1**. The main culture was always inoculated with an appropriate volume of a pre-culture, which was grown overnight at 37 °C, to reach a start OD_{600} of 0.1. This main culture was grown to exponential phase (OD_{600} 0.6 – 0.8) at 37 °C and shifted to optimal expression temperature. Gene expression was induced by adding IPTG to a final concentration of 0.25 – 1 mM. Cells were harvested after 2 – 8 h induction by centrifugation (6000 g, 10 min, 4 °C) and stored at -20 °C until further processing. The expression conditions are summarised in **Table 3-11**.

Table 3-11: Expression conditions.

Gene product	Host	[IPTG]	Induction temperature	Induction time
*Ec*CsgA, CsgA CM, CsgA CD	T7 Express/DL39	1 mM	37 °C	2 h
CsgA – barnase/sumo fusion proteins	BL21 (DE3)	0.25 mM	20 °C	4 h
*Fg*HET-$s_{218-289}$	BL21 (DE3), TUNER(DE3)	1 mM	37 °C	8 h

3.4.3 Production of Isotope Labelled Protein

To obtain uniform isotope labelled protein, gene expression was carried out in M9 medium, which was enriched with an isotope labelled carbon and nitrogen source. Pre-cultures were grown in LB (*Ec*CsgA) or in M9 (*Fg*HET-s) and transferred into the M9 main culture. In case of *Ec*CsgA, the pre-culture was harvested by centrifugation (6000 g, 10 min, 4 °C) and re-suspended in M9 medium. 2 L main culture were incubated until reaching an OD_{600} of 1.0, spun down as described and transferred again in 0.5 L fresh M9 enriched with $^{15}NH_4Cl$ as well as ^{13}C-glucose for double labelling followed by induction with 1 mM IPTG (Marley *et al.*, 2001). To gain isotope labelled *Fg*HET-s the M9 pre-

culture was directly used to inoculate the main culture of M9 supplemented with $^{15}NH_4Cl$ and ^{13}C-glucose. Main culture was grown and induced as shown in **chapter 3.4.2** avoiding centrifugation steps in between.
Growing host bacteria on a defined medium, which is supplemented with appropriate isotope labelled amino acids, leads to protein labelling at specific residues (Muchmore *et al.*, 1989). In order to obtain *Ec*CsgA$_{21-151}$ that is ^{15}N and ^{13}C labelled at positions of aspartatic acid, asparagine, glycine and phenylalanine an auxotrophic *E. coli* strain DL39 was used as expression host (LeMaster and Richards, 1988). Pre- and main cultures were grown in MMM9, whereas the main culture was enriched with the corresponding labelled amino acids and the remaining amino acids in unlabelled form.

3.4.4 Cell Lysis

Cell pellets were thawed on ice and 1 g cells was suspended in 5 mL lysis buffer I supplemented with 1 mg lysozyme per 1 g cell pellet. After incubation for 1 h at 4 °C cells were disrupted by sonication for 10 min on ice (1 s pulse, 5 s delay, sonotrode TT13/F2, Bandelin/Sonopuls). The additional incubation in presence of 5 µg/mL DNase and 5 mM $MgSO_4$ for 30 min at 4 °C ensured DNA degradation. In case of Bar-CsgA-Bar and Sumo-CsgA-Sumo fusion proteins, lysis occurred chemically without sonication. Therefore, 1 g of cells was suspended in 5 mL of lysis buffer II enriched with lysozyme and DNase/$MgSO_4$ as described above. The cell suspension was incubated for 1 h at 4 °C.
Finally soluble and insoluble components were separated by centrifugation (37000 g, 45 min, 4 °C). Except for Bar-CsgA-Bar and Sumo-CsgA-Sumo fusion proteins all proteins were obtained in the insoluble fraction as inclusion bodies.

3.4.5 Purification under Denaturing Conditions and Fibrillisation

Since most of the proteins, which were analysed during this study, were obtained as inclusion bodies after expression and cell lysis, they were purified under denaturing conditions comprising different conditions summarised in **Table 3-12** and illustrated below. The composition of the corresponding buffers is listed in **Table 3-10**.

Table 3-12: Purification under denaturing conditions.

Protein	IB purification	Solubilisation	IMAC	Desalting
*Ec*CsgA$_{21-151}$, *Ec*CsgA$_{40-151}$, *Ec*CsgA$_{21-151}$ XnC	-	8 M GdnHCl → Sol. buffer A	Ni-Sepharose	Sephadex G25
CsgA CM	N-Lauroylsarcosine → IB wash buffer A	8 M GdnHCl → Sol. buffer A	Ni-Sepharose	Sephadex G25
Bar-(CsgA)$_2$-Bar Sumo-(CsgA)$_2$-Sumo, CsgA CD	N-Lauroylsarcosine → IB wash buffer A	Formic acid, 8 M GdnHCl → Formic acid and sol. buffer A	Ni-Sepharose	Sephadex G25
*Ec*CsgB$_{21-151}$	-	8 M GdnHCl → Sol. buffer A	Ni-Sepharose	Sephadex G25

Protein	IB purification	Solubilisation	IMAC	Desalting
*Fg*HET-s$_{218-289}$	Triton X-100 → IB wash buffer B	6 M GdnHCl → Sol. buffer B	Ni-NTA	Lyophilisation and Sephadex G25

Purification and Solubilisation of Inclusion Bodies

Inclusion bodies (IB) of *Ec*CsgA oligomeric variants and *Fg*HET-s were purified and solubilised adapted from a protocol of Rudolph (Rudolph *et al.*, 1997). At first, the IBs were suspended in the corresponding IB wash buffer, which contained Triton X-100 or N-Lauroylsarcosin as detergent. The IB suspension was incubated under mild agitation for 1 h at 4 °C and following spun (37000 g, 30 min, 4° C). To remove further impurities, dissolving in lysis buffer and centrifugation were repeated twice. IBs were solubilised by re-suspending the pellet of the last centrifugation step in the appropriate solubilisation buffer and further incubation at 4 °C under gentle agitation. *Ec*CsgA variants and *Ec*CsgB$_{21-151}$ had to be incubated for 48 h, *Fg*HET-s for 2 h followed by sonication for 10 min on ice (1 s pulse, 5 s delay, sonotrode TT13/F2, Bandelin/Sonopuls). CsgA CD fusion proteins had to be dissolved in 98 % (v/v) formic acid first. Next, the acid was diluted 1/10 by adding solubilisation buffer and pH was adjusted to 7.2 with NaOH. This suspension was incubated for 48 h too. Insoluble material was removed by centrifugation (37000g, 30 min, 4 °C) and solubilised proteins were purified further.

Affinity Chromatography

All proteins in this study carried a hexa histidine-tag and were applied to IMAC. The protein solutions were mixed either with equilibrated Ni-Sepharose or Ni-NTA matrix (Qiagen, 1 mL per 1 L bacterial culture) and incubated for 1 h at 4 °C to ensure binding to the matrix. Subsequently, the solid phase was packed into a column and immobilized proteins were washed threefold with 25 CVs per step (5, 10 and 20 mM imidazole in the corresponding solubilisation buffer). Finally proteins were eluted with 300 mM (3 x 5 CV) and 500 mM (5 CV) imidazole. Amount and purity of protein solutions were determined by SDS-PAGE (see **chapters 3.5.3** and **3.5.4**). Elution fractions of high yield and purity were combined.

Buffer Exchange and Fibrillisation

To allow fibril formation, the processed protein samples were transferred to native buffer lacking any denaturing agent and incubated accordingly.

<u>*Ec*CsgA Variants:</u>
After IMAC the proteinaceous sample were desalted using Sephadex G-25 medium in a Nap-5 or PD-10 column or a self-packed column (GE Healthcare) as follows:
The samples were spun down (20800 g, 10 min, RT), directly loaded on an equilibrated Nap-5 or PD-10 desalting column (GE Healthcare) and desalted as described in the manufactures protocol and summarised in **Table 3-13**. Proteins were eluted with native

buffer A and fibrillised in presence of 2.5 % (w/w) preformed *Ec*CsgA$_{21-151}$ fibrils or 10 % (w/w) preformed *Ec*CsgB$_{21-151}$ fibrils serving as seed.

Table 3-13: Buffer exchange using Nap5 and PD10 columns (GE Healthcare).

Column	1st equilibration	sample volume	2nd equilibration	elution
Nap5	10 mL water, 10 mL buffer	500 µL	-	2 x 400 µL
PD-10	50 mL water, 50 mL buffer	0.5 – 1.5 mL	ad 2.5 mL buffer	4 x 850 µL

Furthermore a self-packed column (Sephadex G-25 medium, 10/40, column volume ~ 30 mL) was used on an ÄKTA-FPLC system to desalt samples. 2 – 3 mL were spun down (20800 g, 10 min, RT) and loaded on the equilibrated column. The proteins were eluted with 1.5 column volumes of native buffer A with a flow rate of 1.5 mL/min.
In the case of cysteine variants, proteins were firstly reduced for 30 min in presence of 20 mM DTT and afterwards processed as described above. The elution buffer contained 5 mM DTT.
Later experiments revealed that a filtration step of the solubilised and processed samples is necessary to obtain concentration dependent lag phases during fibril formation. Thus, proteinaceous samples were filtered before buffer exchange at room temperature by ultracentrifugation (4000 g, RT) using Vivaspin concentrators (MWCO 30000 - 50000). Finally, the permeate was processed further as described above.

*Ec*CsgB$_{21-151}$
Samples of *Ec*CsgB$_{21-151}$ were desalted as described above for the *Ec*CsgA variants. To avoid the premature aggregation of *Ec*CsgB$_{21-151}$, 50 mM KPi pH 3 was used for the desalting procedure. After the elution, the pH was adjusted by adding NaOH and the samples were incubated for fibrillisation.

*Fg*HET-s$_{218-289}$:
From IMAC derived and combined elution fractions of *Fg*HET-s$_{218-289}$ were dialysed (SnakeSkin membrane, 3500 MWCO, Thermo Scientific) twice for 4 h and overnight at 4 °C against 100fold volume of 0.1 % (v/v) acetic acid. Afterwards the protein solution was frozen in liquid nitrogen and lyophilised in order to concentrate the protein sample. The obtained dry powder was dissolved in 0.5 – 1.5 mL of 50 mM Tris/HCl pH 8, 150 mM NaCl, 6 M GdnHCl. Buffer exchange to native buffer B was performed by desalting using Nap-5 columns as described above (**Table 3-13**). The eluted protein was incubated at room temperature under gentle shaking.

Preparation of Seeds

To ensure specific fibrillisation, protein samples of *Ec*CsgA variants were supplemented with preformed fibrils as seeds. These fibrils were incubated up to four weeks (at least one week) at room temperature and prepared accordingly: Fibrils were sedimented (20800 g,

5 min, RT) and resuspended in native buffer A. The centrifugation and resuspension steps were repeated twice. Finally, the fibrils were sonicated for 10 min in an ultrasonic bath and added to freshly desalted protein samples.

3.4.6 Purification under Native Conditions

Bar-CsgA-Bar and Sumo-CsgA-Sumo were partly obtained in the soluble fraction after cell lysis and could be purified under native conditions. Thus, the soluble cell lysate was directly loaded onto Ni-Sepharose and processed as described above. Retained proteins were washed stepwise with 5, 10 and 20 mM imidazole in 50 mM Tris/HCl pH 7.5, 150 mM NaCl, 5 % (v/v) glycerol and eluted with 300 mM and 500 mM imidazole. Elution fractions were analysed via SDS-PAGE and accordingly combined. The proteins were concentrated by ultracentrifugation (4000 g, 4 °C) using Vivaspin concentrators (MWCO 10000) up to 2 mg/mL. Finally, size exclusion chromatography was performed using a Superdex 16/60 prep grade column (GE Healthcare). 2 mL of a concentrated protein sample were loaded onto the equilibrated column and eluted with 50 mM Tris/HCl pH 7.5, 150 mM NaCl. Pooled samples were stored at 4 °C.

3.5 Protein Analytical Methods

3.5.1 Physico-Chemical Parameters of the Studied Proteins

The physico-chemical parameters of the studied proteins are summed up in **Table 3-14** and were calculated using the programs PROTPARAM (www.expasy.org).

Table 3-14: Physico chemical parameters of the studied proteins.

Protein	Position H6	Number of amino acids	MW (kDa)	pI	ε_{280} (M^{-1} cm^{-1})
*Ec*CsgA$_{21-151}$	C-term.	138	14.05	5.73	11460
*Ec*CsgA$_{21-151}$ XnC	C-term.	138	~ 14	5.73	11460
*Ec*CsgA$_{40-151}$	C-term.	119	12.38	5.59	9970
CsgA CM	N-term.	177	18.25	5.56	14440
CsgA CD	N-term.	290	29.77	4.88	24410
Bar-CsgA-Bar	C-term.	347	37.63	6.55	63830
Bar-(CsgA)$_2$-Bar	C-term.	459	48.98	5.71	72310
Sumo-CsgA-Sumo	N-term.	334	36.54	4.88	12950
Sumo-(CsgA)$_2$-Sumo	N-term.	446	47.89	4.75	21430
*Ec*CsgB$_{22-151}$	C-term.	137	14.65	9.3	8940
*Fg*HET-s$_{218-289}$	C-term.	79	9.04	8.26	6990

X is any amino acid, n corresponding residue number.

3.5.2 Photometric Quantification of Protein Concentration

According to Lambert Beer law, the concentration of purified protein solutions was determined by detecting the absorption at 280 nm A_{280} against buffer. The molar extinction coefficient ε_{280} (**Table 3-14**) for each protein was calculated *in silico* (**chapter 3.5.1**).
Under denaturing conditions A_{280} was measured in spectroscopic pure guanidine hydrochloride. Under native conditions absorption was detected immediately after desalting avoiding fibrillisation.

3.5.3 NaDOC/TCA Precipitation

In preparation for SDS-polyacrylamide gel electrophoresis proteins from guanidine hydrochloride containing solutions had to be precipitated by sodium deoxycholate and trichloracetic acid (Arnold and Ulbrich-Hofmann, 1999; Bensadoun and Weinstein, 1976). Corresponding samples were diluted 1:10 in 0.1 % (w/v) NaDOC and mixed with TCA to a final concentration of 10 % (w/v). After incubation for 30 min at 4 °C the samples were spun down (20800 g, 30 min, 4 °C). The obtained pellets were washed twice by adding 500 µL chilled acetone, which was removed after centrifugation (20800 g, 10 min) and the pellets were air-dried to remove retained acetone. Dried pellets were dissolved in 20 µL SDS sample buffer and applied on a SDS gel.

3.5.4 SDS-Polyacrylamide Gel Electrophoresis

In order to analyse amount, purity and molecular weight of protein solutions, discontinuous SDS-polyacrylamide gel (SDS PAGE) electrophoresis was performed under denaturing and reducing conditions (Laemmli, 1970). Therefore, proteinaceous samples were mixed with an excess of SDS-sample buffer, boiled for 5 – 10 min at 95 °C and run on a suitable gel. The gels consisted of a 5 % (w/v) acrylamide stacking gel and a 12 or 15 % (w/v) acrylamide separating gel. Corresponding to the separating gel the PAGE was performed at 40 mA per gel for 28 and 33 min respectively. Furthermore commercial NuPAGE® Novex® 4-12% Bis-Tris Midi Gels (Invitrogen, 1.0 mm x 26 well) were run as described in the manufacturers' manual. If not stated, otherwise always 10 µL were loaded per lane. Finally, gels were stained with Coomassie R-250 and destained with corresponding solutions by boiling and shaking.

3.5.5 Transfer of Proteins to Membranes (Western Blot) and Immunodetection

For immunodetection or N-terminal sequencing experiments protein samples in SDS-polyacrylamid gels were immobilised on polyvinylidene difluoride (PVDF) membranes (Immobilon P, Millipore). First, the membrane was activated in 100 % methanol (~ 10 s) and washed with transfer buffer. Freshly run gels were equilibrated in transfer buffer as well (15 min). Following, the gel was placed onto the membrane, and

both placed between two layers of soaked Whatman paper onto the anode of a semi-dry blot apparatus. The transfer was carried out for 40 min at 15 V.
After western blotting unspecific binding sites on the PVDF membrane were saturated for 1 h in blocking solution and washed in TBST (5 min). Antibodies were diluted in TBST/0.5 % (v/v) blocking solution. The membrane was incubated with the primary antibody (**Table 3-2** in **chapter 3.1.1**) overnight and after washing with TBST (3 x 5 min) additionally incubated for 2 h with the secondary antibody (**Table 3-3** in **chapter 3.1.1**). Finally, the membrane was washed (TBST, 3 x 5 min) and immobilised proteins were detected by BCIP/NBT (Promega) or Enhanced Chemiluminescence (ECL, LumiLight, Roche).

3.5.6 N-Terminal Sequencing

Proteins, which were transferred to a membrane as described in chapter 3.5.5, were excised after visualising with Ponceau S staining solution and sequenced by automated Edman degradation (Edman and Begg, 1967). N-terminal sequencing was carried out by Rita Getzlaff or Beate Jaschok-Kentner (HZI, Braunschweig) using a 494A HT Protein Sequencer (Applied Biosystems).

3.5.7 Mass Spectrometry

MALDI-TOF-MS was used to verify the mass and intactness of recombinant produced proteins. Mass shifts were analysed to confirm site-directed amino acid exchange as well as the incorporation of thiol-active probes. The fibrils (approximately 25 - 50 µg) were dissolved in 98 % (v/v) FA, loaded onto a prespotted anchor chip (PAC) target with an α-cyano-4-hydroxycinnamic acid matrix and dried at RT. To avoid N-formylation the FA containing solutions were processed as fast as possible. The molecular masses were determined in the positive-ion mode on a Bruker Ultraflex time-of-flight mass spectrometer (Bruker Daltonics GmbH). All experiments were performed by Manfred Nimtz, Anja Meier and Undine Felgenträger (HZI, Braunschweig).

3.6 Protein Structure Analysis

3.6.1 Secondary Structure Analysis

CD experiments and FT-IR spectroscopy experiments were performed to probe the secondary structure content of a protein. In CD experiments the differential absorption of left and right circular polarised light by chiral molecules is recorded. FT-IR is based on measuring molecular bond vibrational frequencies. Thus, both techniques provide information about the content of β-sheets, α-helices and loops of a protein structure.

Circular Dichroism

In preparation of CD experiments, protein samples were first transferred to a CD suitable buffer (50 mM KPi pH 7.2 or water) by dialysis of soluble proteins or washing of specific aggregates. If not stated otherwise, protein concentration was adjusted to 1 mM peptide bonds and samples were applied to cuvettes with a path length of 1 mm (Hellma). Spectra were taken using a spectropolarimeter JASCO J-810 and accumulated (5 – 20 spectra). Measurements were performed in far UV (190 – 260 nm) range in stepwise scanning mode with a resolution of 0.2 – 0.5 nm at 4 °C (soluble proteins) or 20 °C (fibrils).

Fourier Transform Infrared Spectroscopy

FT-IR spectroscopy was carried out using a Bruker Tensor 27 system (Bruker Optics) equipped with a BioATR cell (ZnSe single crystal), a mid-IR source (4000 – 400 cm^{-1}) and a KBr beamsplitter. The protein samples were applied as fine suspension with a protein concentration of 1 – 2 mg/mL. In preparation of the FT-IR experiments, fibrils were washed three times with water by repeating centrifugation (20800 g, 5 min, RT) and re-suspension steps. Finally fibrils were re-suspended and sonicated for 10 min in an ultrasonic bath. Measurements were carried out in absorbance mode using following conditions: sample and background scan time was 200 scans, optical filter was open and aperture was set to 6 mm, detector setting was given by LN-MCT Photovoltaic and preamp gain A.

3.6.2 Probing the Amyloid fold as Tertiary/Quaternary Structure

The specific aggregation into the amyloid fold was confirmed by different imaging, staining and biochemical methods.
Since specific aggregates of amyloid fibrils are obtained as long, unbranched fibrous structures, electron microscopy allows to differ between unspecific and specific aggregates.
Most amyloids can be stained by the amyloid dyes CR and ThT. The binding of CR is detected as green birefringence under cross-polarised light. The interaction of ThT with amyloid fibrils results in an increased and robust ThT fluorescence at 485 nm.
Furthermore, the use of limited proteolysis was used to identify stable folded protein domains. Thus, this technique was also used to probe amyloids stability.

Electron Microscopy

In order to analyse fibrils by electron microscopy, appropriate samples were deposited on a carbon-coated grid and stained negatively by a method adapted from Valentine *et al.*, 1968. Appropriate samples were negatively stained with 4% (w/v) aqueous uranyl acetate using 300 mesh nickel grids (Plano). To coat these grids, thin carbon support films, approximately 10 nm thick, were prepared. Carbon was sublimated from a carbon thread on to freshly cleaved mica (SCD500, Bal-Tec). Finally, the negatively stained samples

were examined in a Zeiss 910 transmission electron microscope (Zeiss) with an acceleration voltage of 80 kV and at calibrated magnifications. Images were recorded digitally with a Slow-Scan CCD-Camera (ProScan, 1024x1024) with ITEM-Software (Olympus Soft Imaging Solutions). All experiments were performed by Manfred Rohde and Heinrich Lühnsdorf (HZI, Braunschweig)

Congo Red Labelling and Birefringence

In order to stain specific amyloid aggregates with congo (Westermark *et al.*, 1999), 20 µL of fibril suspension were air-dried on a glass slide and adjacent incubated with 10 µM congo red (CR stock solution diluted in water) until the dye was also dried. Excess of CR was washed away by rinsing with 70 % (v/v) ethanol. Typical birefringence was detected under cross polarised light using a microscope with two polarisation filters.

Thioflavin-T Fluorescence Studies

If not stated otherwise, fluorescence measurements were carried out under stirring using a Fluoro-Max4 spectrometer (Horiba Jobin Yvon) and corresponding fluorescence cells with a light path of 10 x 4 mm (Hellma). Steady-state and time-resolved Thioflavin-T fluorescence studies were carried out to detect protein structural changes (fibril formation) over time.
For steady-state experiments protein solutions (2 – 50 µM) were mixed with a two-fold molar excess of ThT and incubated for 30 min to allow dye binding. Samples were excited at 442 nm (band width 5 nm) and emission was followed from 465 to 600 nm (band width 5 nm). The ThT fluorescence was correlated to the intrinsic fluorescence of the protein (excitation 280 nm, emission 340 nm).
In time-resolved ThT studies 200 µL of fresh desalted CsgA variants (2 – 16 µM) were applied to a 96 well plate (Corning, Flat Bottom, Non-Binding Surface, Non-Sterile, Black Polystyrene) and subsequently incubated in presence of 20 µM ThT. ThT fluorescence intensity was obtained at 485 nm after excitation at 442 nm every 10 min using a Tecan-Platereader. The measurements were performed at 25 °C. For both, excitation and emission, a band width of 5 nm was used and multiple readings were obtained per well (4 x 4, border 1000 µM). The homogeneity of the samples was ensured by shaking the plate before every measurement for 2 s (amplitude 2 mm). The gain (100 – 150) and the z-position (~ 20000 µm) were adjusted manually before starting the assay using preformed fibrils.

Limited Proteolysis

In order to compare the stability of different CsgA variants, the obtained aggregates were applied to limited proteolysis. First *Ec*CsgA fibrils were washed twice by re-suspending in proteolysis buffer and centrifugation (20800 g, 5 min, RT). 35 µg fibrils were incubated at room temperature with proteinase K in different ratios (1:5 – 1:1000 (w/v)). Proteolytic digestion was stopped by adding the 4-fold volume of 98 % (v/v) FA and the acid was

removed using a speed vac at 60 °C for 1 h. Remained pellets were dissolved in 20 µL urea/SDS solution and analysed via SDS-PAGE.

3.6.3 Analysis of Side-Chain Solvent Accessibility and Inter- and Intramolecular Contacts

Due to their natural properties, β-sheets are always characterised by a specific architecture. The side-chains of neighbouring amino acids in a β-strand face to opposite directions. Thus, amyloid fibrils are characterised by solvent accessible and not accessible side chains. Depending on the arrangement of the β-strands within the fibril, amyloid structures feature specific intra- and intermolecular contacts. As cysteine residues can be labelled with thiol-reactive agents, they are sensitive probes for solvent accessibility and inter- and intramolecular contacts (Javitch *et al.*, 2002). These thiol-reactive probes comprise for example sodium iodacetate (IAA), iodacetamide (IAM), and the fluorescence dyes AlexaFluor$_{488}$ C_5 maleimide (AF488) and N-(1-pyrene)maleimide (PM) summarised in **Table 3-15** (Molecular Probes, Invitrogen; Hansen and Winther, 2009).

Table 3-15: Thiol-reactive agents and their properties.

Agent	Molecular mass (g mol^{-1})	Absorption (nm)	Emission (nm)	Detection
IAA	186	-	-	Mass spectrometry
IAM	185	-	-	Mass spectrometry
AF488	643	~ 495	~ 520	Mass spectrometry, Fluorescence spectroscopy
PM	297	~ 345	450 – 500 nm (excimer)	Mass spectrometry (labelling), Fluorescence spectroscopy (excimer)

Since these agents bind covalently to the thiol groups, the labelling efficiency can be determined by mass or in case of fluorescence dyes by an increased fluorescence signal. Furthermore pyrene-labelling can lead to the formation of excimers that are detected by a strong red shift in fluorescence (Lehrer, 1997). This excimer formation only occurs when two cysteine bound pyrene molecules lie within 4 – 10 Å. Thus, inter- and intramolecular contacts can be analysed.

Specific Thiol-Labelling

In order to analyse the β-strand topology as well as inter- and intramolecular contacts within *Ec*CsgA fibrils, IAA, IAM, AF488 and PM (Molecular Probes, Invitrogen) were applied to appropriate *Ec*CsgA$_{21\text{-}151}$ cysteine variants. To optimise the labelling strategy, several different labelling conditions (**Table App-1**) were checked. Two main strategies were followed: labelling of fibrils under native conditions (post assembly) and labelling of monomeric, unfolded protein under denaturing conditions (pre assembly).

One post assembly labelling strategy was identified as optimal and used solely: The appropriate cysteine variants were fibrillised as described (**chapter 3.4.5**). Prior desalting, proteins were supplemented with 10 % (w/w) solubilised *Ec*CsgA$_{21-151}$ wt to enable co-aggregation into the wt fold. To seed the fibrillisation, 2.5 % (w/w) prefomed wt fibrils were added. After two days incubation under mild agitation, the observed fibrils were subjected to labelling experiments adapted to former protocols (Krishnan and Lindquist, 2005; Ritter *et al.*, 2005). In preparation of labelling, the fibrils were washed twice with labelling buffer (50 mM Tris/HCl pH 8, 150 mM NaCl, 0.1 mM TCEP) by repetitive centrifugation (20800 g, 5 min, RT) and resuspension. 20 - 100 µM of two days old fibrils were incubated in presence of an appropriate thiol-reactive agent for 30 min (2.5 molar excess AF488) and 3 h (20x IAA, 20x IAM, 10x PM) respectively. The sample volume was around 250 - 500 µl. The labelling reaction was quenched by centrifugation (20800 g, 5 min, RT) and three washing steps comprising repetitive re-suspension in 50 mM KPi pH 7.2 and centrifugation (20800 g, 5 min, RT). In the first washing step the buffer was supplemented with 5 mM DTT.

Detection of Specific Thiol-Labelling

The chemical labelling of thiol groups was detected via mass spectrometry and fluorescence spectroscopy respectively.
In case of mass spectrometry analysis, fibrils were harvested by centrifugation as described above and solubilised in 98 % (v/v) FA. After solubilisation the samples were processed as fast as possible to avoid N-formylation as described in **chapter 3.5.7**.
To analyse the linkage to AF488 and excimer formation of PM, the fibrils were re-suspended in 1 mL 50 mM KPi pH 7.2 and applied to a Fluoro-Max spectrometer (Horiba Jobin Yvon) in a fluorescence cell with a light path of 10 x 4 mm (Hellma). AF488 fluorescence was detected at 516 nm (1 nm band width) after excitation at 493 nm (1 nm band width). The observed values were normalised to the intrinsic fluorescence (excitation 280 nm, emission 340 nm). PM excimer fluorescence was analysed in a range of 360 to 600 nm (2 nm band width) after excitation at 343 nm (3 nm band width). An emission maximum around 450 – 500 nm was characteristic for excimer fluorescence (Lehrer, 1995).

3.6.4 NMR spectroscopy

General Measurement Conditions

All NMR experiments were carried out at 25 °C using a Bruker Avance III 600 spectrometer equipped with four radio-frequency channels and a 5 mm Z-axis gradient triple-resonance cryo-probehead. The offset that represses the water signal and the pulse length were optimised for every individual dataset.
Amyloid fibrils were collected by centrifugation (20800 g, 4 min, RT) and solubilised in ~ 500 µL aprotic DMSO (perdeuterated, d_6-DMSO) containing up to 0.1 % (v/v) TFA (deuterated, d_1-TFA), and the samples were placed in a 5 mm NMR tube (Norrel). For 1D-

and 2D-experiments ~ 20 mg wet pellet of fibrils was processed, for 3D-experiments up to 35 mg were used.
Under native conditions 50 mM KPi pH 6 was used as buffer and complemented with 10 % (v/v) D_2O prior measurements. The protein concentration was ~ 50 µM.
All experiments and most important parameters are summed up in **Table 3-16**, sample preparation is shown in **Table 3-17**. The spectra were processed and analysed using the programs PROSA (Guntert *et al.*, 1992) and CARA (Keller, 2004).

Table 3-16: NMR experiments carried out in this work and their parameters. SW: sweep width (spectral width), TD: time domain (number of obtained data points), NS: number of scans

Experiment	Obtained correlation	SW (Hz)	TD	NS	pulse programme
		1D-experiments			
1D-^{1}H	^{1}H	8417.5	2048	128	zgpr
		2D-experiments			
fHMQC (Luhrs *et al.*, 2005; Ritter *et al.*, 2005)	^{1}H(i), ^{15}N(i)	4807.7 1338.3	1024 128	2	fhmqc, time optimised HMQC (modified from Bax *et al.*, 1983)
ctHSQC	^{1}H(i), ^{15}N(i)	4807.7 1338.3	2048 192	2	NHSQC-ct (Mori *et al.*, 1995)
		3D-experiments			
NOESY-HSQC	^{1}H(i)⟶ ^{1}H(j), ^{15}N(j)	7812.5 1338.3 7803.7	4096 128 320	8	noesy.ct-hsqc.G (Marion *et al.*, 1989)
HNH-NOESY	^{1}H(i)⟶ ^{1}H(j), ^{15}N(j)	8417.5 8403.4 1338.3	2048 384 72	8	HNH_noesy (Marion *et al.*, 1989)
HNCA	$^{1}H^{N}$(i), ^{15}N(i), $^{13}C_{\alpha}$(i), $^{13}C_{\alpha}$(i-1)	8417.5 1338.3 4830.6	4096 62 196	16	hncapwg3d (Grzesiek and Bax, 1992)
HN(CO)CA	$^{1}H^{N}$(i), ^{15}N(i), $^{13}C_{\alpha}$(i-1)	8417.5 1338.3 4830.4	4096 62 196	16	hncocapwg3d (Grzesiek and Bax, 1992)
HNCACB	$^{1}H^{N}$(i), ^{15}N(i), $^{13}C_{\alpha}$(i), $^{13}C_{\alpha}$(i-1), $^{13}C_{\beta}$(i), $^{13}C_{\beta}$(i-1)	8417.5 1338.3 11312.2	2048 64 192	32	hncacbgp3d (Wittekind and Mueller, 1993)

Table 3-17: Sample preparation for NMR experiments.

Protein	Experiment	wet pellet	[TFA]	volume
*Ec*CsgA	1D & 2D	20 – 25 mg	0.1 % (v/v)	520 µL
	3D	35 mg		
*Fg*HET-$s_{218-289}$	1D & 2D	~ 20 mg	0.05 % (v/v)	500 µL
	3D	20 mg		

Sequence Specific Resonance Backbone Assignment

^{13}C,^{15}N-labelled fibrils were subjected to triple-resonance-experiments HNCACB (Wittekind and Mueller, 1993), HNCA and HN(CO)CA (Bax and Ikura, 1991) as well as HNH nuclear Overhauser enhancement spectroscopy (Diercks *et al.*, 1999) experiments to achieve the sequence-specific resonance assignment of the backbone H^N/N cross-peaks. HNCACB, HNCA and HN(CO)CA enabled the correlation of backbone amide proton and nitrogen resonances with the corresponding alpha- and beta-carbon resonances. These experiments resolve the proton/nitrogen correlations in the same fashion as [^{15}N,^{1}H] correlation experiments and the third dimension of the spectra contains ^{13}C-chemical shifts of carbon resonances summarised in **Table 3-16** and **Figure 3-2**.

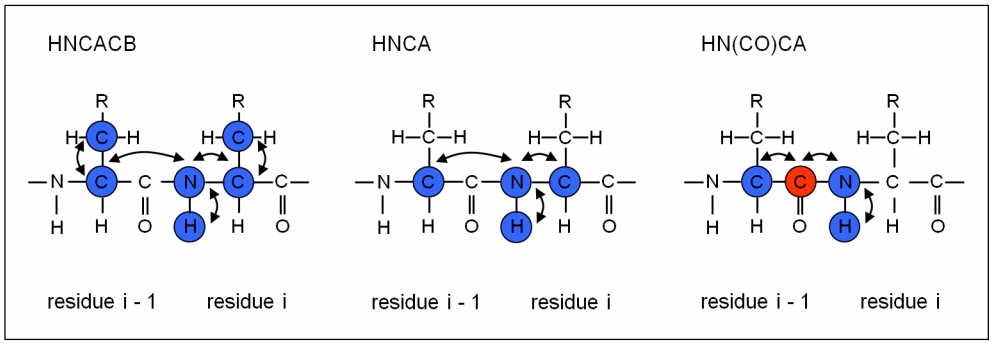

Figure 3-2: Magnetisation transfer during the triple-resonance experiments HNCACB, HNCA and HN(CO)CA. The nuclei, whose frequencies are detected, are indicated in blue. The nuclei, which only transfer the magnetisation, are shown in red. The arrows indicate the magnetisation direction.

Due to different magnetisation transfer pathways, HNCACB spectra display ^{13}C-chemical shifts of the C_α and C_β resonances of the same residue as the backbone amide and the preceding residue in one strip. HNCA spectra display C_α shifts of a given residue and the predecessor, while HN(CO)CA spectra exploit the C_α shift of the previous residue only.
C_α- and C_β-resonances are differentiated by their opposite sign. Resonances of the same residue as the backbone amide and the preceding residue are discriminated by their intensities. Based on this information, peaks are aligned establishing that the corresponding amide pairs are adjacent in the sequence. The C_α- and C_β-chemical shifts were compared to the average of chemical shifts for each amino acid deposited in the BMRB (www.bmrb.wisc.edu). This allowed the assignment of the strips onto the protein sequence. The assignments were confirmed by HNH nuclear Overhauser enhancement spectroscopy experiments. Each strip (3rd dimension) of a NEOSY-HSQC spectrum contains NOEs from one backbone amide group to all protons close by. Due to solubilising the fibrils, proteins were unstructured and the obtained signals corresponded only to short and medium range NOEs enabling the alignment of adjacent backbone amides.

H/D Exchange NMR Experiments

^{15}N labelled fibrils of *Ec*CsgA and *Fg*HET-s variants were used for H/D exchange studies of the backbone amides (Hoshino *et al.*, 2002; Li and Woodward, 1999).

To start the exchange reaction, the fibrils were sedimented at 20800 g for 4 min, washed in corresponding native buffer comprising D_2O as the solvent, sedimented again and re-suspended in the same buffer for incubation up to 12 weeks at room temperature. Every sample contained around 20 – 25 mg wet pellet of fibrils. Hydrogen exchange was quenched at suitable intervals (15 min, 1 h, 1 d, ..., 20 w) by sedimenting the fibrils (20800 g, 4 min, RT) and freezing the pellet in liquid nitrogen.
A series of up to 80 two-dimensional [^{15}N,^{1}H] correlation spectra were recorded directly after solubilising the fibrils in ~500 µL perdeuterated dimethyldisulphoxide (d_6-DMSO) containing deuterated trifluoric acid (d_1-TFA). The amount of residual D_2O was about 4 % (wet pellet of fibrils). Residues that display a fast exchange in the fibrils and residues with high intrinsic exchange rates in DMSO result in absent peaks in the [^{15}N,^{1}H] correlation spectrum. To differentiate them, a second series of two-dimensional spectra were measured after addition of 4 % (v/v) H_2O.
Next, the peak volumes within a series of NMR spectra were integrated for each amino acid and extrapolated to t_{DMSO} = 0 min. The extrapolated peak volume displayed the resonance intensities corresponding to the incubation time of fibrils in D_2O (Equation 3-1). To determine the specific exchange rates these intensities were fitted monoexponential to the exchange time (**Equation 3-1**).

$$I_t = \left(I_0 + I_\infty\right) \cdot e^{-k_{ex} \cdot t} + I_\infty$$

Equation 3-1: Exponential decay. I_t gives peak volume at certain time, I_0 peak volume at t = 0, I_∞ at t_∞, k_{ex} exchange rate and t certain time.

The data were analysed by using the programs PROSA (Guntert *et al.*, 1992) and CARA (Keller, 2004), and a special written program in Visual basic in combination with Microsoft Excel (Luhrs *et al.*, 2005).

4 Results

4.1 Choice of Experimental Strategy

As outlined in the introduction (**chapter 1.1**) and in the aims and scope (**chapter 2**), the two functional amyloids *Fg*HET-s$_{218-289}$ and the major curli subunit *Ec*CsgA were addressed for structural analysis.

Since amyloids form filamentous structures that are inherently difficult to study, conventional methods for protein-structure analyses are usually not feasible. During this study two strategies (**Figure 4-1**) were followed to investigate the structure of amyloids and amyloid-like intermediates. Thus, fibrils of *Fg*HET-s$_{218-289}$ and *Ec*CsgA variants, as well as probable soluble monomeric analogues of *Ec*CsgA oligomers were addressed to study.

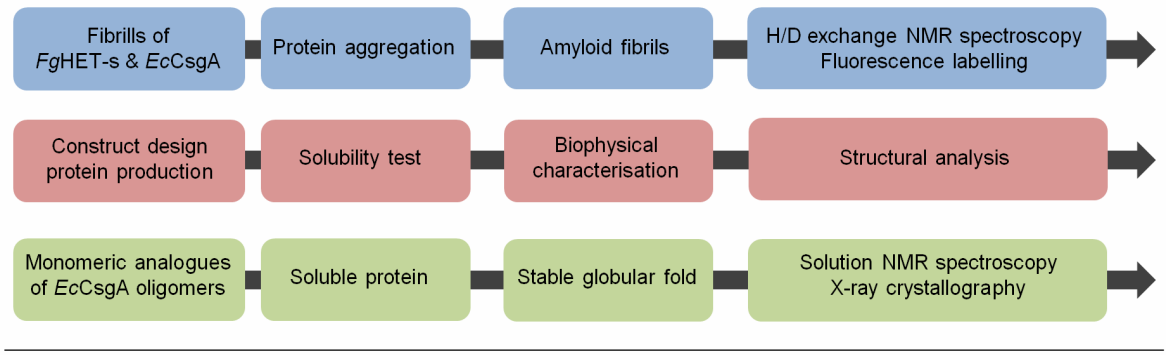

Figure 4-1: Experimental strategy.

The recent developments in biophysical methods facilitate closer insights into fibrillar structures. X-ray crystallography, as well as H/D exchange NMR and solid state NMR spectroscopy enable sequence specific structural information of amyloids (reviewed by Chiti and Dobson, 2006; Toyama and Weissman, 2011). As illustrated in **chapter 1.4.2**, X-ray crystallography is only suitable for the analysis of short peptide fragments forming amyloid-like crystals. But H/D exchange NMR and solid state NMR spectroscopy facilitate the structural investigation of wt full-length amyloid fibrils (**chapter 1.4.4**, **1.4.5** and citations therein).

H/D exchange NMR spectroscopy was chosen as main technique in this study to analyse the fibril topology of *Fg*HET-s$_{218-289}$ and *Ec*CsgA even though solid state NMR spectroscopy has the potential for the amyloid structure elucidation at atomic resolution.

To gain atomic insights into protein structures by solid state NMR spectroscopy, numerous experiments and distance measurements are necessary. Although with *Pa*HET-s$_{218-289}$ the first amyloid structure was recently solved by solid state NMR experiments (Wasmer *et al.*, 2008), for most proteins uniformly labelled experiments are often less helpful due to peak overlapping. The crowded spectra hinder precise assignments and distance measurements.

H/D exchange NMR spectroscopy is a powerful tool to identify structured regions in amyloids at single residue resolution. Furthermore, the determination of H/D exchange rates provides information about the stability of amyloids as well as the homogeneity and regularity within the fibrils (Luhrs *et al.*, 2005; Ritter *et al.*, 2005; Vilar *et al.*, 2008). Thus, H/D exchange NMR spectroscopy is highly suitable to compare homologous proteins and their variants. In combination with other techniques like fluorescence labelling studies, H/D exchange NMR provides valuable insights into the fibril topology and was chosen for the current questions about the fibril topology of *Fg*HET-s$_{218-289}$ and *Ec*CsgA.

In order to obtain soluble monomeric analogues of *Ec*CsgA oligomers that are suitable to study their stable soluble structure, the aggregation tendency of the amyloidogenic *Ec*CsgA was to be blocked. Due to their natural properties, β-sheets are always characterised by edge strands, which exhibit free hydrogen bond donors or acceptors. Thus, these aggregation-prone β-strands can always template the interaction with any other β-strands by hydrogen bonding. Natural soluble β-sheet proteins are protected by different features against these interactions. Due to the missing improvement of the β-sheet structure in general, this natural strategy is termed negative design. The strategy comprises a) covering edge strands with a loop of a helical or nonrepetitive structure, b) inward rolling of one sheet, c) the introduction of a local twist and a local outward protrusion by a β bulge and/or a proline and d) an inward pointing charged side chain (Richardson and Richardson, 2002; Cherny *et al.*, 2005; Soto *et al.*, 1998; Wang and Hecht, 2002). Since proline and charged residues are easily introduced by site-directed mutagenesis, the strategies of implementing a local twist by proline and inward pointing charged side chains were to be checked for *Ec*CsgA. Furthermore, the N- and C-terminal fusion of small well folding proteins was chosen to cover *Ec*CsgA edge strands against β-sheet interactions. Potentially, soluble and stable folded proteins are suitable for conventional structural analysis.

4.2 Structural Characterisation of the *Fg*HET-s Prion Domain

In order to explain the observed cross-seeding between two HET-s homologues, *Fg*HET-s$_{218-289}$ was subjected to H/D exchange NMR experiments to compare its structure with *Pa*HET-s. This analysis was done as part of a structure-function study of the HET-s homologues in collaboration with the labs of Beat Meier (ETH Zürich) and Sven Saupe (Université de Bordeaux). Christian Wasmer (ETH Zürich) kindly provided the recombinant plasmid encoding *Fg*HET-s$_{218-289}$.

4.2.1 Protein Production and Purification

The HET-s prion forming domain of *Fusarium graminearum Fg*HET-s$_{218-289}$ was purified as C-terminal histidine-tagged protein after overexpression in *E. coli* BL21 (DE3) cells as described in **chapter 3.4.5** and illustrated in **Figure 4-2**.

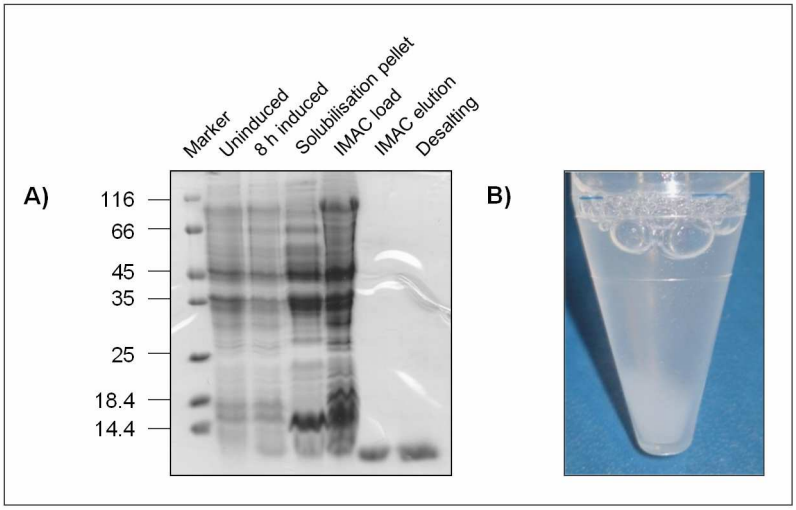

Figure 4-2: Purification and aggregation of *Fg*HET-s$_{218-289}$. (A) 15 % SDS-PAGE (Coomassie staining) of *Fg*HET-s$_{218-289}$ production, isolation by IMAC and desalting. **(B)** *Fg*HET-s$_{218-289}$ aggregates obtained after 1 d incubation.

At first, cultivation and cell harvest were followed by cell lysis and IB purification via several resuspension and centrifugation steps. After solubilisation the protein solution was subjected to Ni-NTA to perform IMAC under denaturing conditions. Protein containing fractions were pooled, dialysed against 1 % (v/v) acetic acid and afterwards lyophilised to concentrate the protein sample. The dry powder was re-suspended in 6 M GdnHCl and the protein solution was desalted using Sephadex G25. Desalting resulted in monomeric

*Fg*HET-s$_{218-289}$, which started immediately to aggregate similar to *Pa*HET-s$_{218-289}$ (Balguerie *et al.*, 2003; Ritter *et al.*, 2005).

Thus, *Fg*HET-s$_{218-289}$ was purified up to homogeneity and polymerised. The aggregation was completed within a few hours.

4.2.2 Recombinant *Fg*HET-s$_{218-289}$ Forms Amyloid Fibrils *in Vitro*

Prior to the structural investigation, *Fg*HET-s$_{218-289}$ was demonstrated to fibrillise specifically. To ensure the analysis of amyloid fibrils, the observed aggregates were characterised by CD, CR binding, ThT fluorescence and EM as shown in **Figure 4-3**.

Far-UV-CD spectra of soluble and insoluble material showed a typical shift of the ellipticity (Ranjbar and Gill, 2009). This shift from a minimum at 200 nm to a minimum at 218 nm and a maximum below 200 nm indicated the conversion from the random coil conformation to a β-sheet rich fold similar to what has been shown for *Pa*HET-s$_{218-289}$ (Balguerie *et al.*, 2003).

CR binding resulted in green birefringence under cross polarised light as expected for amyloids (Hawe *et al.*, 2008).

In contrast to *Pa*HET-s$_{218-289}$ (Sabate *et al.*, 2007), ThT mixing experiments with *Fg*HET-s$_{218-289}$ aggregates at neutral pH resulted in a strong ThT fluorescence. ThT is suggested to bind to channels within the β-sheets of the amyloid fibrils. Since ThT is positively charged, basic residues hamper the ThT binding (Krebs *et al.*, 2005). In the bundles of *Pa*HET-s$_{218-289}$ fibrils all channels feature at least one positively charged, basic residue that probably inhibits ThT binding (Wasmer *et al.*, 2008). In comparison to *Pa*HET-s$_{218-289}$, *Fg*HET-s$_{218-289}$ features a different charge distribution as indicated above in the sequence alignment in **Figure 1-6**. In case of a similar fold, the channels within *Fg*HET-s$_{218-289}$ fibrils would possess less basic residues. This charge distribution probably enables an increased ThT binding.

Finally, bundles of laterally associated fibrils were obtained by EM similar to *Pa*HET-s$_{218-289}$ (Balguerie *et al.*, 2003).

In summary, the biophysical characterisation of the observed aggregates verified specific amyloid fibril formation by *Fg*HET-s$_{218-289}$ *in vitro*. Thus, the aggregates were capable for specific H/D exchange NMR experiments.

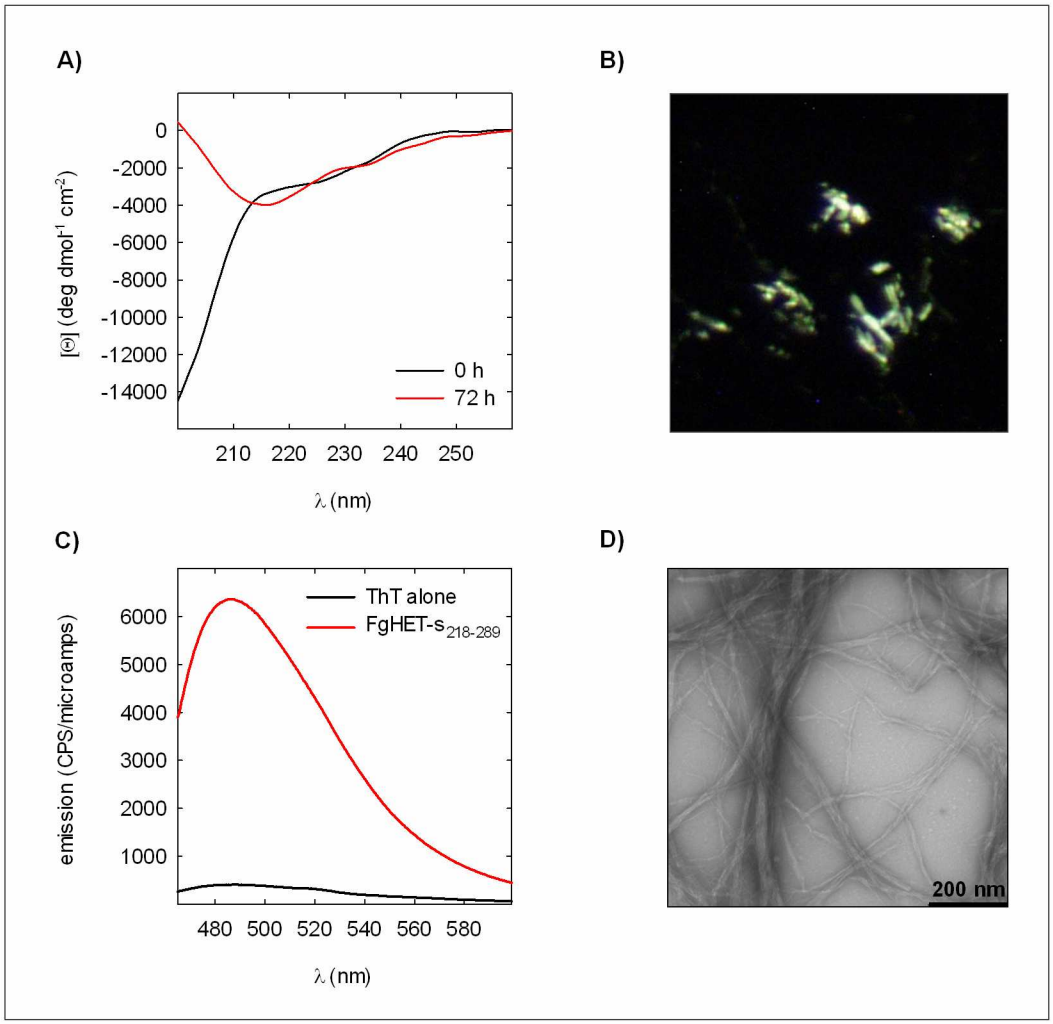

Figure 4-3: Biophysical characterisation of *Fg*HET-s$_{218-289}$. (A) Far-UV-CD spectra of soluble *Fg*HET-s$_{218-289}$ in 50 mM Tris/HCl pH 8, 150 mM NaCl directly after desalting (black) and aggregates after 3 d incubation (red). 20 spectra were accumulated using a cuvette with 1 mm path length. **(B)** Micrograph of *Fg*HET-s$_{218-289}$ fibrils in presence of 10 µM aqueous CR showing green birefringence under cross polarised light. **(C)** ThT fluorescence in presence of *Fg*HET-s$_{218-289}$ (red) compared to buffer (black, 50 mM Tris/HCl pH 8, 150 mM NaCl). 40 µM ThT were excited at λ_{ex} = 442 nm in presence and absence of 20 µM protein. **(D)** Electron micrograph of *Fg*HET-s$_{218-289}$ aggregates.

4.2.3 Sequence-Specific Secondary Structure Analysis

For the structural comparison of the two cross-seeding HET-s homologues, regularly structured segments in *Fg*HET-s$_{218-289}$ fibrils were specified. This analysis included the sequence-specific backbone-resonance assignment of *Fg*HET-s$_{218-289}$ and appropriate H/D exchange NMR experiments of specific *Fg*HET-s$_{218-289}$ fibrils.

Sequence-Specific Backbone-Resonance Assignment

In order to analyse the H/D exchange behaviour of each backbone amide, the sequence specific resonance assignments of the backbone amides were obtained by triple-resonance HNCACB experiments (Wittekind and Mueller, 1993) and HNH nuclear Overhauser enhancement spectroscopy (Diercks *et al.*, 1999) using uniform ^{13}C,^{15}N labelled *Fg*HET-s$_{218-289}$. Based on corresponding ^{13}C chemical shifts and NOE signals each backbone amide was assigned to one cross-peak with the exception of R252 (**Figure 4-4**).

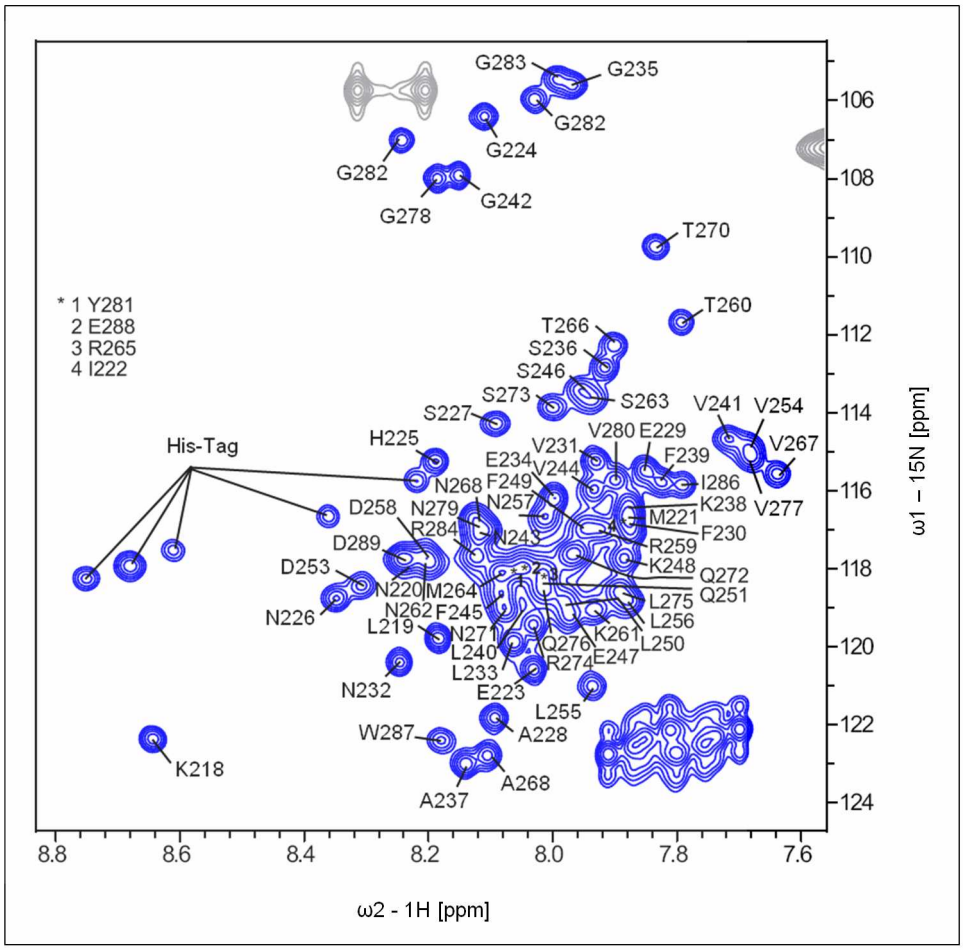

Figure 4-4: Backbone NH-assignments of *Fg*HET-s$_{218-289}$. Fast HMQC spectrum of uniformly ^{15}N labelled *Fg*HET-s$_{218-289}$ (20 mg wet pellet of fibrils) in d_6-DMSO containing 0.1 % (v/v) d_1-TFA. Sequence-specific chemical shift assignments are indicated next to the resonance.

H/D Exchange NMR and Location of β-sheets

Solvent-protected protons of secondary structural elements were identified by hydrogen exchange NMR experiments similar to former studies of amyloid fibrils (Luhrs *et al.*, 2005; Ritter *et al.*, 2005). As described in **chapter 3.6.4**, ^{15}N labelled *Fg*HET-s$_{218-289}$ fibrils were incubated in D_2O buffer up to 12 weeks and H/D exchange was quenched at suitable time points (0 h, 1 h, 24 h, 72 h, 1 w, 2 w, 4 w, 12 w). Subsequently, series of two-dimensional [^{15}N,^{1}H] correlation spectra were recorded by performing fast HMQC experiments. In general, the [^{15}N,^{1}H] fHMQC experiments resulted in spectra with broad, intense peaks after a short acquisition time (~ 2.5 min), which favours a high time resolution and kinetic analyses as outlined in **chapter 1.4.4**. Due to less overlap of the cross-peaks in the [^{15}N,^{1}H] fHMQC spectra of labelled *Fg*HET-s$_{218-289}$, the broad peaks were unambiguously evaluated. Thus, the observed spectra enabled an accurate analysis of the backbone resonances as illustrated below.

The hydrogen exchange of *Fg*HET-s$_{218-289}$ resulted in NMR spectra with different resonance intensities depending on the incubation time in D_2O buffer. After 4 weeks, the intensity of about 40 % of the resonances had disappeared. This suggested that the corresponding amides have undergone exchange with solvent deuterons, which are not visible in the [^{15}N,^{1}H] NMR spectrum (**Figure 4-5**).

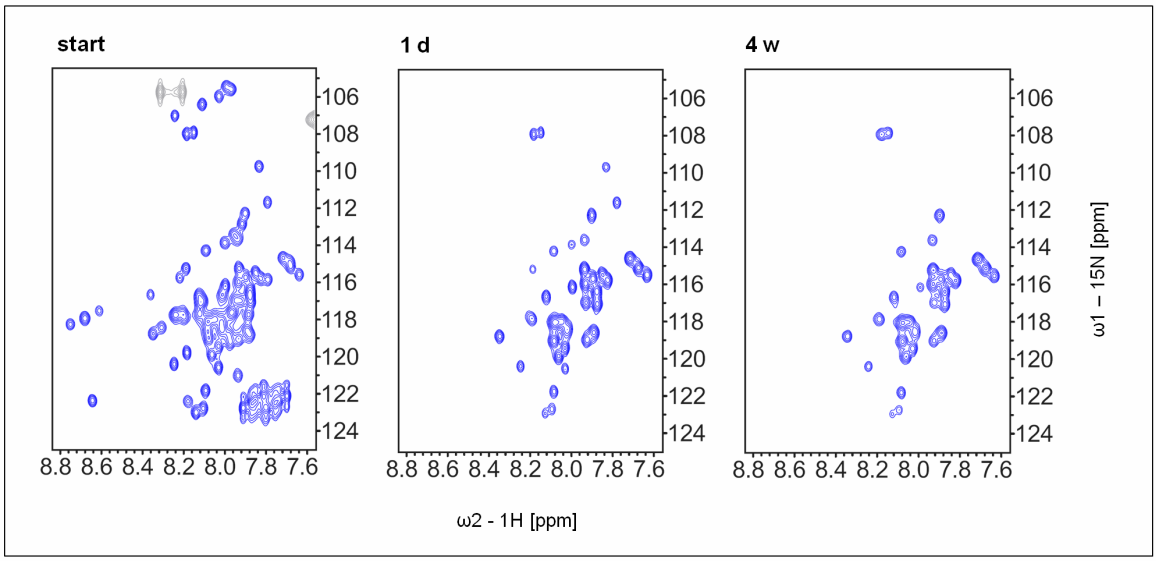

Figure 4-5: H/D exchange of *Fg*HET-s$_{218-189}$. Fast HMQC spectra of uniformly ^{15}N labelled *Fg*HET-s$_{218-289}$ (~20 mg wet pellet of fibrils) in d6-DMSO containing 0.05 % (v/v) d_1-TFA corresponding to fully protonated (start) and partially hydrogen exchanged (1 h, 24 h, 4 w) *Fg*HET-s$_{218-189}$ fibrils.

Subsequently to the addition of H_2O, a second series of [^{15}N,^{1}H] NMR spectra revealed that there are no artefacts of disappearing resonances due to ultra fast H/D exchange in

DMSO. Fast intrinsic exchange in DMSO was still detectable due to the high time resolution of the series of fHMQC experiments (**Figure 4-6 A**).

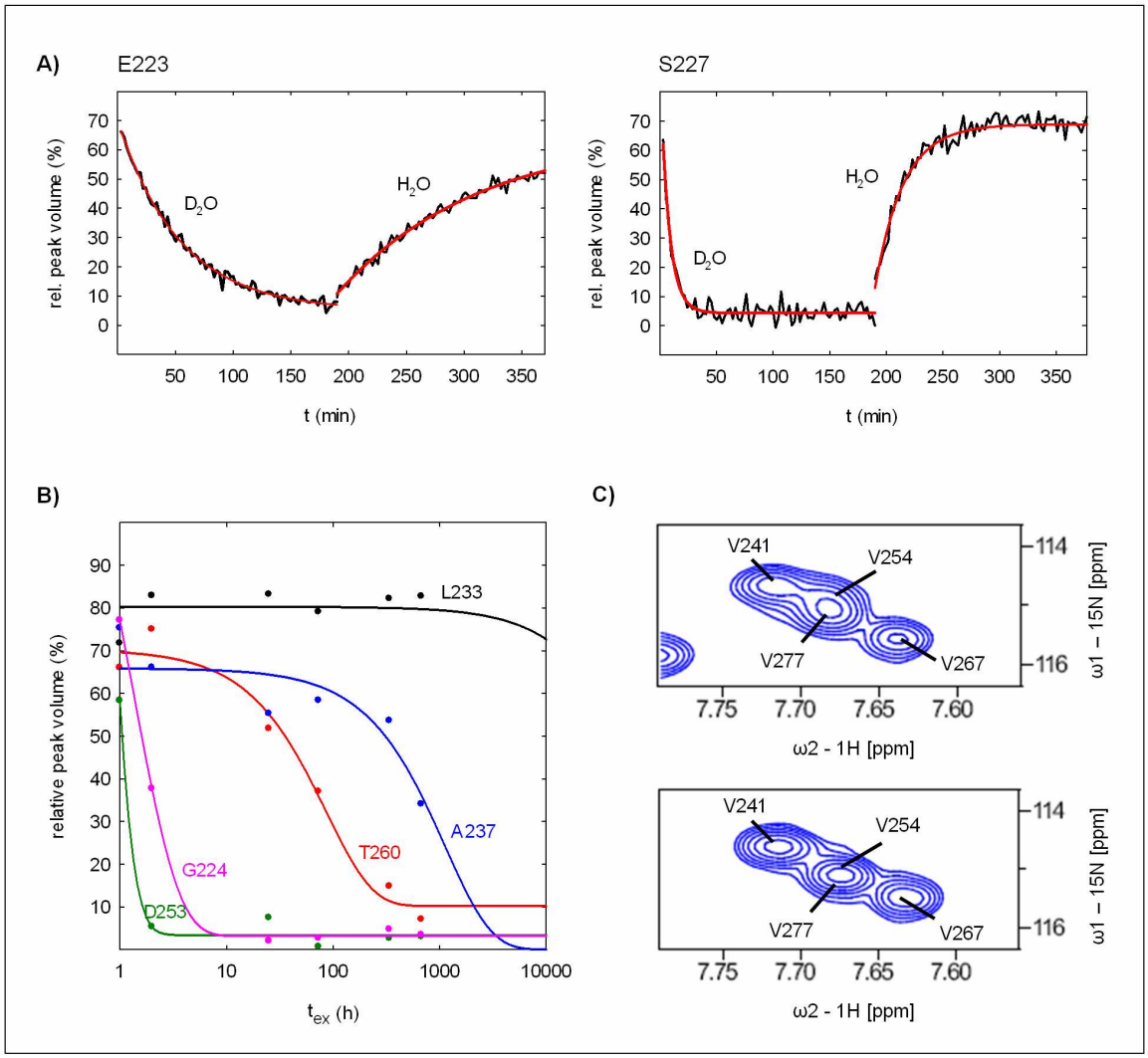

Figure 4-6: Determination of exchange rates. (A) Intrinsic H/D exchange in DMSO of E223 and S227 backbone amides indicated by the changes of peak volumes after monomerising *Fg*HET-s$_{218-289}$ fibrils in DMSO/TFA and final addition of H_2O respectively. Red curves show the monoexponential fits of the raw data. **(B)** Monoexponential fit of peak volumes at t_{DMSO} = 0 min to t_{ex} for selected backbone amides. **(C)** Fast HMQC spectra of uniformly ^{15}N labelled *Fg*HET-s$_{218-289}$ (~20 mg wet pellet of fibrils) in d6-DMSO containing 0.05 % (v/v) d_1-TFA corresponding to fully protonated (0h) and partially hydrogen exchanged (1 d) *Fg*HET-s$_{218-189}$ fibrils indicated strong overlap in the shown region.

Thus, the absent resonances related to backbone amides, which were rapidly substituted by deuterons during incubation of the fibrils in D_2O buffer and not by remained traces of D_2O while recording the NMR spectra. Hence, the peak volumes at t_{DMSO} = 0 min were determined for all assigned resonances and used to calculate accurate exchange rates (**Figure 4-6 B**). Due to strong overlap with resonances, which corresponded to protected backbone amides, residue V254 was excluded from further analysis (**Figure 4-6 C**).

The resulting exchange rates are shown in **Figure 4-7**. All residues revealed a monophasic exchange behaviour indicating a well defined and homogeneous structure of the fibrils. The backbone amides of the five amino-terminal residues, the seven carboxy-terminal residues and residues 246-258 exchanged quickly (≥ 1.5 h^{-1}). They are only weak or not protected against the solvent. Therefore, these residues seem to be conformationally disordered. Four segments were identified displaying slow exchange rates in the range of 10^{-2} h^{-1} to 10^{-5} h^{-1} similar to *Pa*HET-$s_{218-289}$ (Ritter *et al.*, 2005) and $A\beta_{1-42}$ fibrils (Luhrs *et al.*, 2005). These slow exchange rates demonstrate the extraordinary stability and compactness of the studied fibrils. Comprised residues E223-E234, A237-F245, R259-T270 and S273-G282 are protected and thus considered to be involved in hydrogen bonds. Since circular dichroism spectra of the *Fg*HET-$s_{218-289}$ fibrils confirmed a β-sheet rich fold, the four protected segments were identified to form β-strands. The corresponding residues are located in the same regions as described for *Pa*HET-$s_{218-289}$ (Ritter *et al.*, 2005; Wasmer *et al.*, 2008). In contrast to *Pa*HET-$s_{218-289}$, *Fg*HET-$s_{218-289}$ features six more protected backbones comprising E223-H225 and R259-K261.

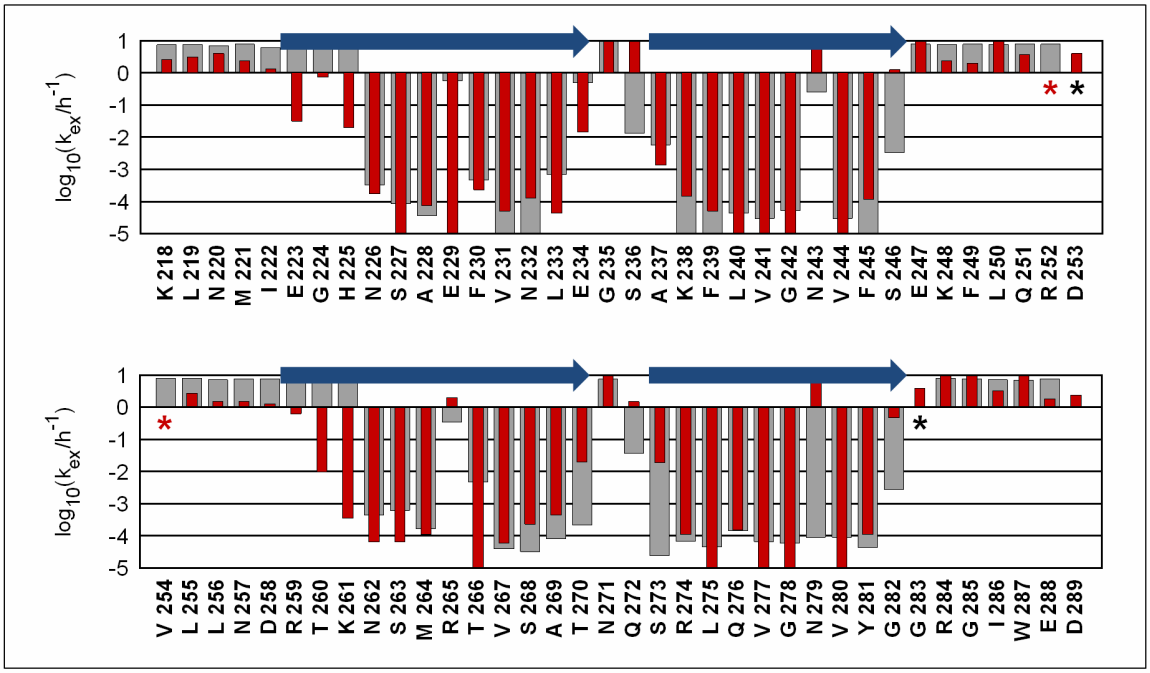

Figure 4-7: H/D exchange data for *Fg*HET-$s_{218-289}$ compared to *Pa*HET-$s_{218-289}$ and secondary structure prediction for *Fg*HET-$s_{218-289}$. Red bars show the exchange rates of *Fg*HET-$s_{218-289}$ k_{ex} (h^{-1}) and blue arrows indicate β-strands. Grey bars indicate rates of *Pa*HET-$s_{218-289}$ (Ritter *et al.*, 2005). The shown amino acid sequence corresponds to that of *Fg*HET-$s_{218-289}$. Residues marked with an asterisk were excluded from the analysis due to missing assignment and overlap respectively.

The residues G235 and S236 as well as N271 and Q272 exchanged fast – indicating β-turns, which connect the segments 1-2 and 3-4 respectively. The 2nd, 3rd and 4th segments

were characterised by outstanding residues: N243, R265 and N279, which displayed a fast exchange rate in a protected region representing a structural disturbance such as a β-arc. E229, N243, R265 and N279 are located at positions, which were identified as less defined in *Pa*HET-$s_{218-289}$ due to forming β-arcs (Wasmer *et al.*, 2008). In contrast to N243, R365 and N279, E229, positioned at a β-arc too, was highly protected against H/D exchange within *Fg*HET-$s_{218-289}$ fibrils. The expected partner R265 revealed fast H/D exchange indicating different hydrogen bonding patterns in the β-arcs at these positions. A similar H/D exchange behaviour was observed in *Pa*HET-$s_{218-289}$ at β-arc positions N243 and N279. N279 was fully protected, while N243 revealed relative fast H/D exchange (Ritter *et al.*, 2005).

Similar to the residues of the β-turns and β-arcs, protected amide backbones of segment 1 and 3 as well as 2 and 4 are separated by 36 residues supporting a repetitive structural pattern.

Despite little differences, the H/D exchange NMR analysis of *Fg*HET-$s_{218-289}$ revealed a structural pattern that is highly similar to that of its homologues *Pa*HET-$s_{218-289}$. The comparison with complementary data of *Pa*HET-$s_{218-289}$ enabled already further insights into the topology of *Fg*HET-$s_{218-289}$ fibrils as discussed in **chapter 5.1**.

4.3 Structural Characterisation of *Ec*CsgA Fibrils

4.3.1 Starting Point of *Ec*CsgA Structure Elucidation

At the beginning of this study an initial H/D exchange analysis of mature *Ec*CsgA (residues 21-151) revealed overcrowded low resolution fHMQC spectra (Alexander Eberth, unpublished). The strong overlap of numerous cross-peaks hampered the initial peak picking and the sequence specific backbone assignment. Moreover, the strong overlapping effect inhibited the precise determination of the resonance intensities after numerous incubation times in D_2O. Thus, the exchange rates were masked and this fact disabled an accurate H/D exchange analysis. For that reason a strategy was developed to improve the spectra and the resonance assignment of *Ec*CsgA comprising an optimised sample preparation to avoid heterogeneity, the application of another NMR experiment, new *Ec*CsgA constructs and selective labelling. Based on the developed concept, H/D exchange experiments were successfully performed and analysed. Furthermore the topology analysis comprised fluorescence labelling studies to elucidate solvent accessible side chains, as well as intra- and intermolecular contacts.

4.3.2 Sequence Analysis and Design of New *Ec*CsgA Constructs

In order to obtain CsgA variants that are suitable for an enhanced structural analysis via H/D exchange NMR spectroscopy and fluorescence labelling studies, appropriate gene constructs were designed.

As shown in **Figure 4-8**, the mature form of *Ec*CsgA consists of a glycine-rich, flexible N-terminal region and five repeating units with the consensus sequence $Sx_5QxGxGNxAx_3Qx_{1\text{-}5}$, x is any amino acid (Collinson *et al.*, 1999; Hammar *et al.*, 1996). The repetitive sequence and especially the glycine/asparagine patches led to a strong overlap in NMR spectra and were found to be a drawback in the resonance assignment and the H/D exchange NMR analysis. As the repeating units are supposed to form the amyloid core (Collinson *et al.*, 1999), at first a truncated *Ec*CsgA variant *Ec*$CsgA_{40\text{-}151}$ was designed. The corresponding gene *csg*$A_{40\text{-}151}$ was cloned into a pET-11d vector using *Nco*I and *BamH*I restriction sites.

In order to elucidate the topology of *Ec*CsgA, cysteine variants of *Ec*$CsgA_{21\text{-}151}$ were analysed with thiol-reactive fluorescence dyes (Javitch *et al.*, 2002; Lehrer, 1997). Cysteine residues were introduced by site directed mutagenesis and the corresponding

oligonucleotides are summarised in **Table 3-8**. The substitution of intramolecularly conserved residues was avoided with the exception of S88C, which constitutes a conservative replacement. Even- and odd-numbered residues where chosen for mutagenesis and labelling studies in order to identify solvent accessible side chains suitable for the investigation of intra- and intermolecular contacts. The corresponding gene constructs are listed in **Table 3-7**.

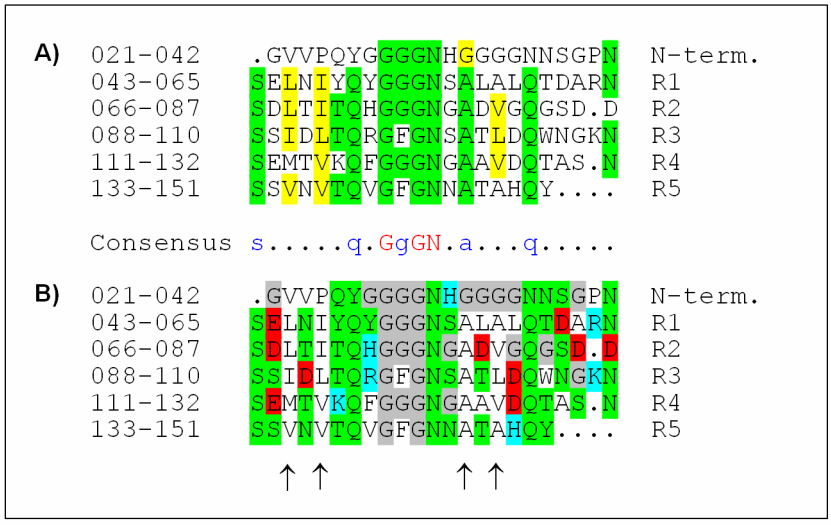

Figure 4-8: Intramolecular alignment of *Ec*CsgA$_{21-151}$. Comparison of the primary structure of the mature N-terminus and the five repeating units. **(A)** Intramolecularly conserved residues are shown in green, homologous residues in yellow. The consensus sequence indicates intramolecularly conserved residues in red and residues conserved only among the repeating units in blue. The analysis was done with the program MultAlin (Corpet, 1988). **(B)** Charged residues are highlighted in red or blue, polar uncharged residues are shown in green, glycines in grey and unpolar residues are not coloured. Arrows indicate hypothetical hydrophobic ladders.

4.3.3 Protein Production and Purification

In contrast to former studies (Wang *et al.*, 2007), *Ec*CsgA variants were purified as C-terminal histidine tagged proteins from inclusion bodies after overexpression in *E. coli* T7 Express cells as described in **chapter 3.4.5** and summarised in **Figure 4-9**. After cultivation and cell harvest, cells were lysed under native conditions and the insoluble material was solubilised subsequently. The solubilisation was followed by IMAC under denaturing conditions using Ni-Sepharose. Eluted proteins were combined and finally desalted using Sephadex G25. To ensure complete buffer exchange conductivity was monitored. Only protein fractions devoid of denaturant were pooled and fibrillised.

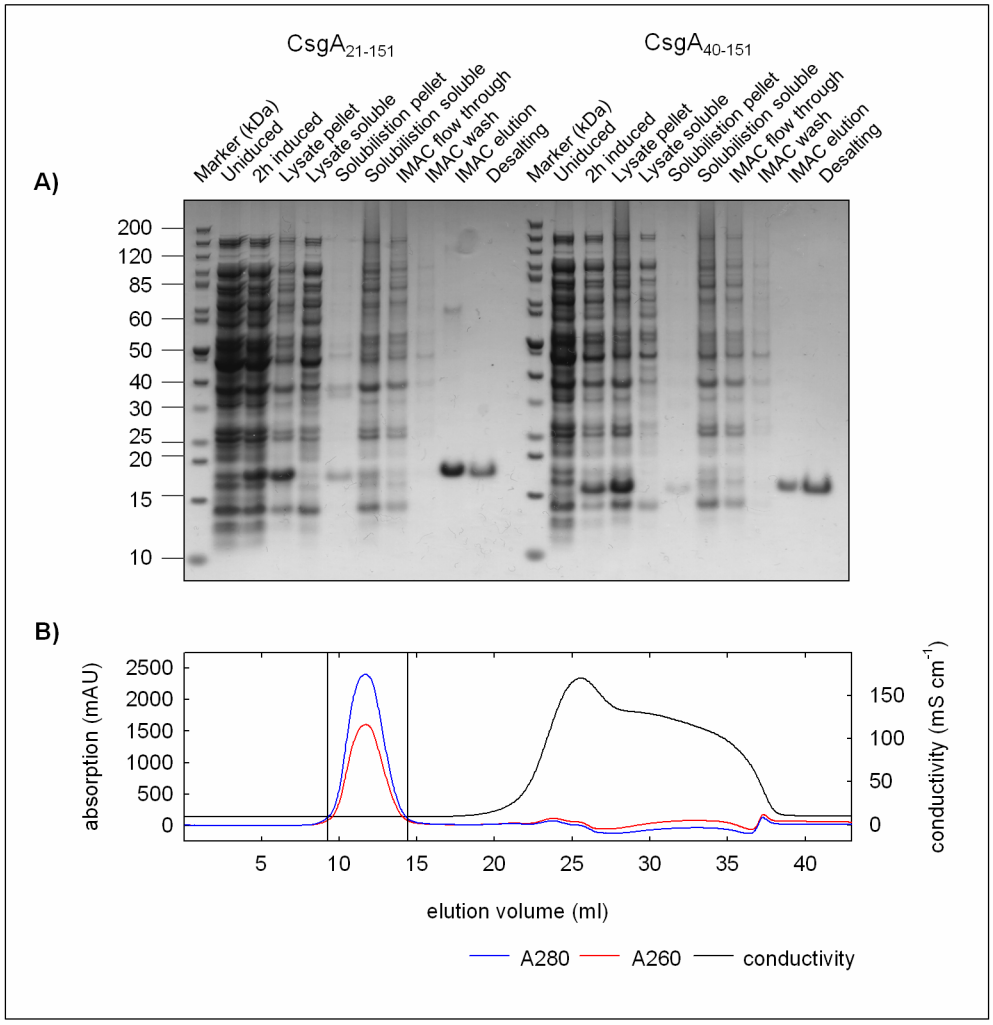

Figure 4-9: Purification of *Ec*CsgA variants. (A) 4-12 % NuPAGE® (Coomassie staining) of production and isolation of *Ec*CsgA$_{21-151}$ and *Ec*CsgA$_{40-151}$ by IMAC and desalting. **(B)** Elution profile of desalting *Ec*CsgA$_{40-151}$ via Sephadex G25 (10/40). Elution occurred via 50 mM KPi pH 7.2 with a flow rate of 1.5 mL/min. The enframed fraction was further characterised.

Thus, *Ec*CsgA variants were purified to homogeneity from bacteria grown in LB as well as minimal medium with a yield of 5 - 10 mg per L culture. Aggregation after buffer exchange was completed within a few hours and the aggregates were stable against 0.1 % (w/v) SDS overnight.

4.3.4 *Ec*CsgA$_{21-151}$ and *Ec*CsgA$_{40-151}$ Have Similar Biophysical Properties

Both variants *Ec*CsgA$_{21-151}$ and *Ec*CsgA$_{40-151}$ readily formed aggregates. FT-IR and CD spectroscopy, as well as ThT binding and EM were used to characterise the amyloid properties of the variants shown in **Figure 4-10**.

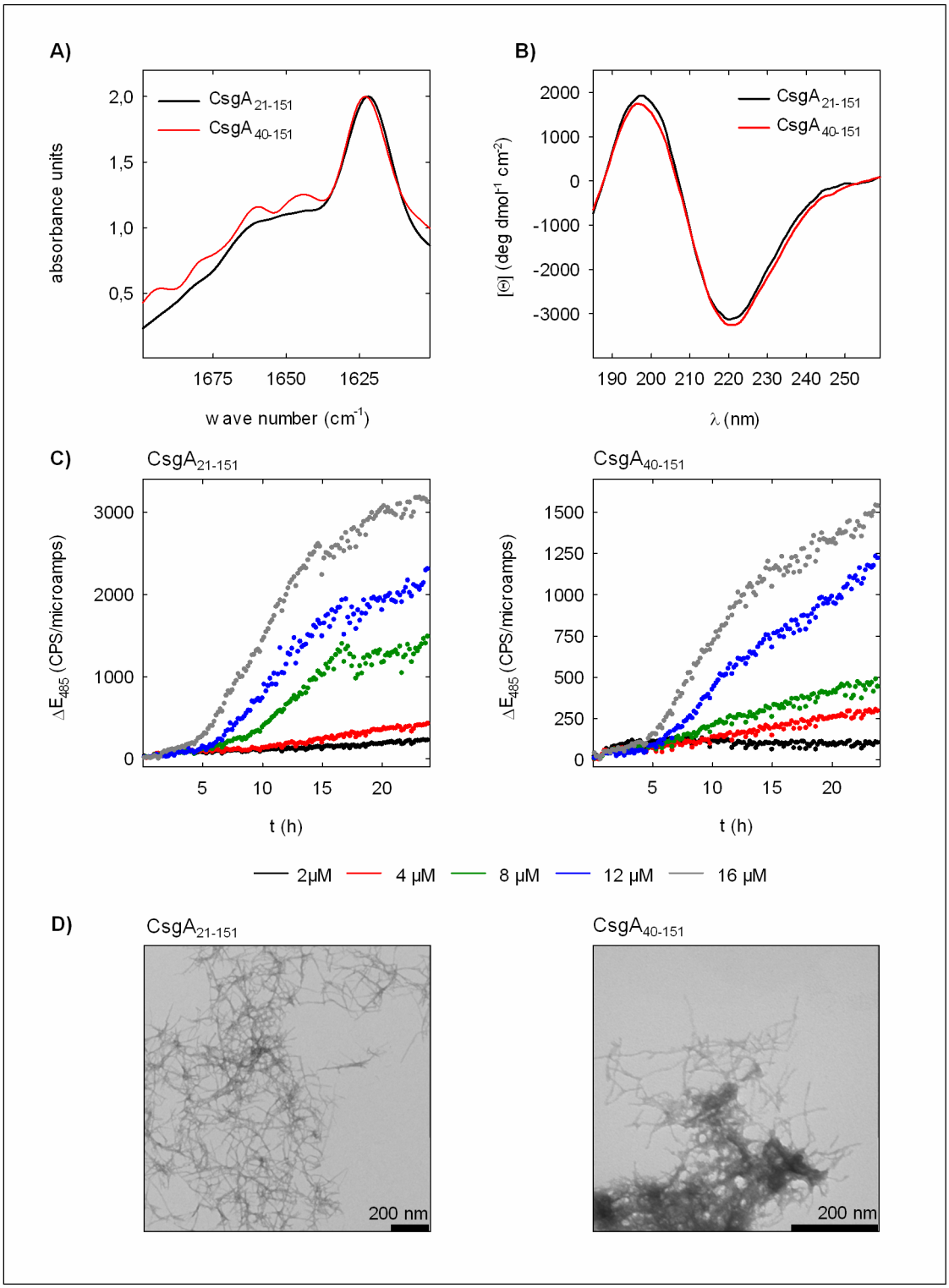

Figure 4-10: Biophysical characterisation of *Ec*CsgA$_{21-151}$ and *Ec*CsgA$_{40-151}$. (A) FT-IR spectra showing the amide I band of aggregates (1 mg/mL) formed by *Ec*CsgA$_{21-151}$ (black) and *Ec*CsgA$_{40-151}$ (red) after 4 days incubation. Aggregates were washed with water and applied to a BioATR cell. **(B)** Far-UV-CD spectra of aggregates in ddH_2O formed by *Ec*CsgA$_{21-151}$ (black) and *Ec*CsgA$_{40-151}$ (red) after 4 days incubation. 20 spectra were accumulated using a cuvette with 1 mm path length. **(C)** Increase in ThT fluorescence depending on time and protein concentration. 25 µM ThT were mixed with freshly desalted *Ec*CsgA using 2 µM (black), 4 µM (red), 8 µM (green), 12 µM (blue) and 16 µM (grey) protein concentration. Fluorescence was measured every 10 min at 485 nm after excitation at 442 nm. **(D)** Electron micrographs of *Ec*CsgA aggregates.

FT-IR and CD were used to estimate the secondary structure content of the aggregates. Both variants displayed a characteristic signal around 1622 cm^{-1} in the FT-IR spectra, which is characteristic for the β-sheet arrangement in amyloid fibrils (Zandomeneghi *et al.*, 2004). CD spectra showed a minimum around 218 nm and maximum around 195 nm, typical for proteins with a high β-sheet content (Ranjbar and Gill, 2009) corresponding to the observations reported for *Ec*$CsgA_{21-151}$ (Chapman *et al.*, 2002; Shewmaker *et al.*, 2009; Wang *et al.*, 2007). Mixing each variant with ThT resulted in a strong ThT fluorescence. ThT was also used to monitor the transition of freshly desalted *Ec*CsgA variants from monomeric proteins into specific amyloid aggregates (LeVine, 1999). For both variants the evolution of the ThT fluorescence showed a distinct lag phase, growth and stationary phase in a protein concentration dependent manner. The observation of concentration dependent lag times was fully consistent with former studies on *Ec*CsgA (Wang *et al.*, 2007) and other amyloidogenic proteins (Lomakin *et al.*, 1997). Electron microscopy revealed regular, unbranched fibrils characteristic for amyloid fibrils as curli (Chapman *et al.*, 2002; Shewmaker *et al.*, 2009; Wang *et al.*, 2007).

In summary, the biophysical analysis confirmed specific aggregation of *Ec*$CsgA_{21-151}$ and *Ec*$CsgA_{40-151}$ into highly similar amyloid fibrils allowing comparative H/D exchange NMR experiments.

4.3.5 Strategies to Improve NMR Spectra and the Assignment

Well resolved and intense NMR spectra are indispensable for a complete resonance backbone assignment and a significant H/D exchange NMR analysis. A first analysis of *Ec*$CsgA_{21-151}$ failed due to insufficient resolved NMR spectra. In order to improve the NMR spectra and the assignment of *Ec*CsgA, another type of [^{15}N, ^{1}H] correlation experiments and new *Ec*CsgA constructs, as well as selective labelling were employed.

ctHSQC Experiments Resulted in Spectra With High Resolution

[^{15}N, ^{1}H] NMR spectra that display well resolved peaks are essential for accurate H/D exchange NMR analysis. As fHMQC experiments of *Ec*$CsgA_{21-151}$ resulted in spectra with high intensity but insufficient resolved peaks, ctHSQC experiments were employed to *Ec*CsgA variants. These experiments differ in the evolution of magnetisation as well as the evolution period. In ctHSQC a constant ^{15}N evolution period is used instead of the

conventional variable evolution period. In the case of *Ec*CsgA$_{21-151}$, ctHSQC experiments resulted in [^{15}N,^{1}H] correlation spectra with high intensity, well separated peaks and less peak broadening as shown in **Figure 4-11**. This enhancement enabled accurate peak picking and the precise determination of resonance intensities for separated peaks.

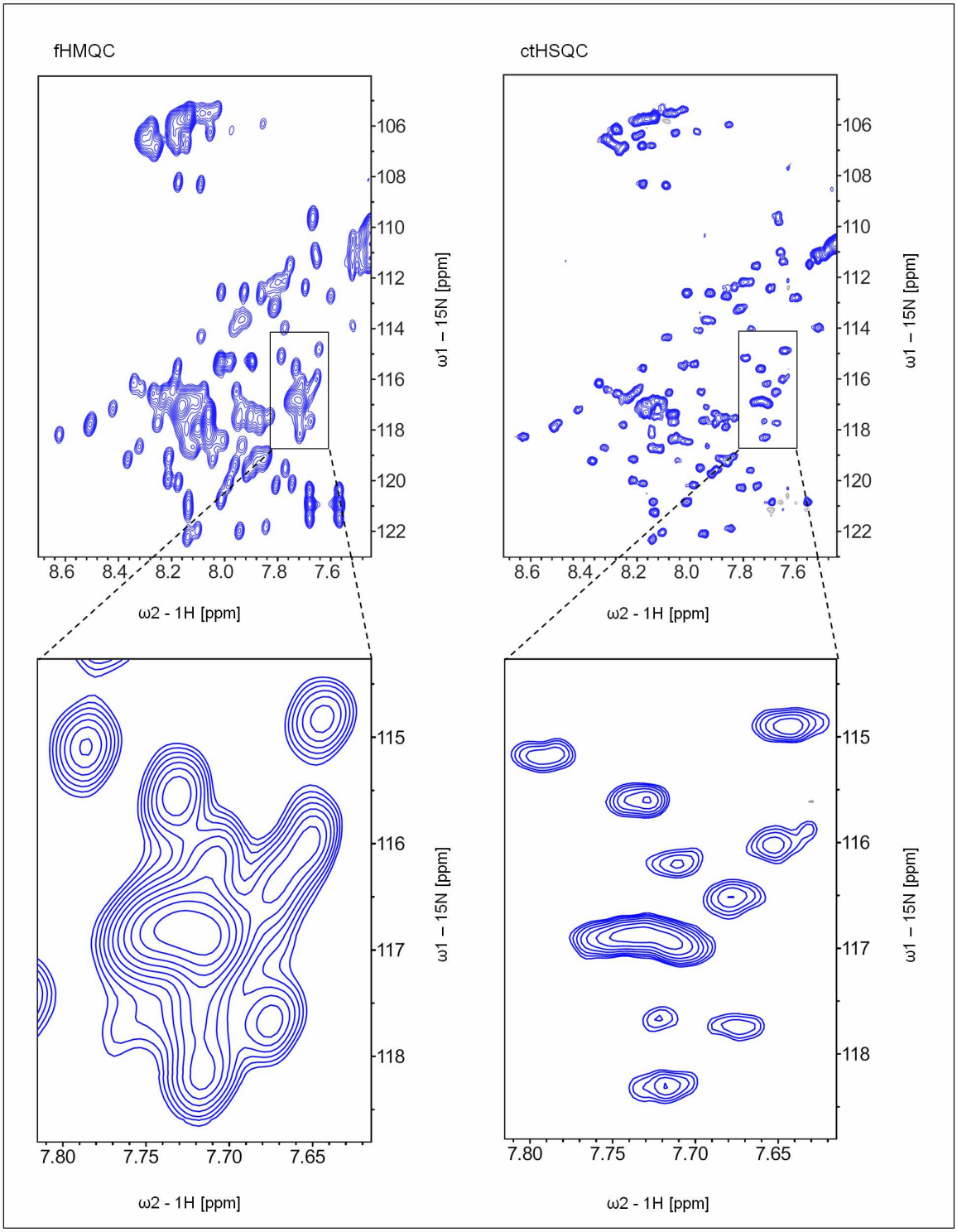

Figure 4-11: fHMQC and ctHSQC spectra of *Ec*CsgA$_{21-151}$ in comparison. fHMQC and ctHSQC spectrum of 20 mg wet pellet of uniformly ^{15}N labelled CsgA$_{21-151}$ fibrils in d_6-DMSO containing 0.1 % (v/v) d_1-TFA.

Improved Spectra of New *Ec*CsgA Variants

In addition to enhanced 2D NMR experiments, new *Ec*CsgA variants and selective labelled samples (**Figure 4-12**) were utilised to promote the *Ec*CsgA assignment.

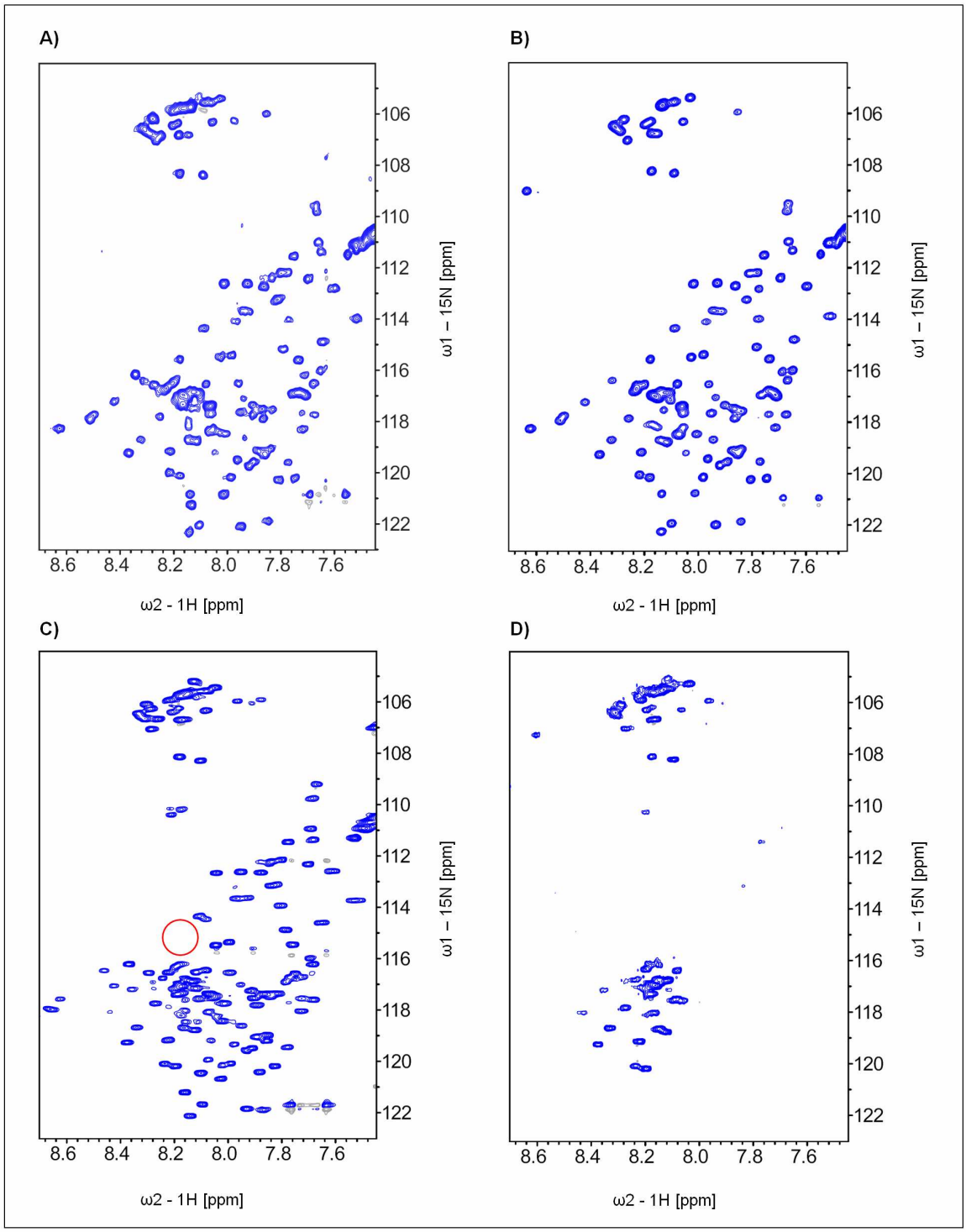

Figure 4-12: ctHSQC spectra of different *Ec*CsgA variants. All spectra were taken from 20 – 25 mg wet pellet of appropriate labelled fibrils, which were solubilised in d_6-DMSO containing 0.1 % (v/v) d_1-TFA. **(A)** *Ec*CsgA$_{21\text{-}151}$. **(B)** *Ec*CsgA$_{41\text{-}151}$. **(C)** *Ec*CsgA$_{21\text{-}151}$ S55C. The red circle corresponds to the original position of the cross-peak of S55. **(D)** *Ec*CsgA$_{21\text{-}151}$ specific labelled at positions of Asn, Asp, Gly and Phe.

The repetitive sequence and especially the glycine/asparagine patches led to a strong overlap in NMR spectra and were found to be a drawback in the resonance assignment and the H/D exchange NMR analysis (Alexander Eberth, unpublished). As the repeating units are supposed to form the amyloid core (Collinson *et al.*, 1999), the truncated *Ec*CsgA variant *Ec*CsgA$_{40\text{-}151}$ was subjected to NMR experiments. The corresponding spectra demonstrated less overlap of the cross-peaks due to lacking the resonances of the N22 tag illustrated in **Figure 4-12 A/B**.

Furthermore, cysteine variants of *Ec*CsgA$_{21\text{-}151}$ were employed. In comparison to *Ec*CsgA wt, the peaks of the substituted amino acid and adjacent residues were dislocated (**Figure 4-12 C**). Owing to the dislocation of the peaks, corresponding assignments were confirmed.

Selective labelling studies are common to support sequence specific assignments. In order to assign the glycine/asparagine patches, an *Ec*CsgA$_{21\text{-}151}$ sample, which was specifically labelled at the positions of glycine, asparagine, aspartatic acid and phenylalanine, was produced and analysed in NMR experiments. The resulting spectra displayed solely the resonances of the labelled residues shown in **Figure 4-12 D**.

4.3.6 Sequence-Specific Analysis of Secondary Structure Elements

In order to elucidate the fold of the amyloid fibrils formed by *Ec*CsgA and to study the effect of the nucleator protein *Ec*CsgB to this fold, sequence-specific positions of regular secondary structure elements in fibrils of *Ec*CsgA$_{21\text{-}151}$ and *Ec*CsgA$_{40\text{-}151}$ as well as heteronucleated *Ec*CsgA$_{40\text{-}151}$ fibrils were analysed. These analyses were done analogously to the precedent study of *Fg*HET-s$_{218\text{-}289}$ (**chapter 4.2.3**). In addition to HNCACB and NOESY experiments, the sequence-specific backbone-resonance assignment of *Ec*CsgA depended on numerous NMR experiments comprising different variants and different labelling strategies. In contrast to the H/D exchange NMR analysis of *Fg*HET-s$_{218\text{-}289}$, [^{15}N,^{1}H] correlation spectra of *Ec*CsgA variants were recorded by performing ctHSQC experiments instead of fHMQC experiments. As described above, ctHSQC experiments resulted in intense and well resolved peaks after ~ 5 min acquisition time enabling an accurate analysis of single spectra (**chapter 4.3.5**). Compared to the short acquisition time of fHMQC experiments (~ 2.5 min), the acquisition time of ctHSQC experiments is less favourable in respect of an appropriate time resolution. Despite the lower time resolution,

series of ctHSQC spectra still enabled the correction of the resonance intensities for exchange in DMSO and thus reasonable H/D exchange NMR analysis as shown below.

Sequence-Specific Backbone-Resonance Assignment

Sequence specific resonance assignments of the backbone amides of *Ec*CsgA were finally obtained by a combination of NMR experiments recorded on different *Ec*CsgA variants employing selective labelling strategies summarised in **Table 4-1**.

Table 4-1: NMR experiments used for *Ec*CsgA sequence specific backbone resonance assignment.

***Ec*CsgA variant**	**Labelling**	**2D experiments**	**3D experiments**
*Ec*CsgA$_{21-151}$	uniform	fHMQC, ctHSQC	HNCACB, HNCA, HNCOCA, HNH NOESY
*Ec*CsgA$_{21-151}$	Asp, Asn, Gly, Phe	fHMQC, ctHSQC	HNCACB
*Ec*CsgA$_{21-151}$ XnC	uniform	ctHSQC	-
*Ec*CsgA$_{40-151}$	uniform	ctHSQC	HNCACB, NOESY-HSQC

X is any amino acid, n corresponding residue number.

Finally 88 % of the backbone amides of *Ec*CsgA$_{21-151}$ and 95 % of *Ec*CsgA$_{40-151}$ were assigned to the sequence illustrated in **Figure 4-13**. In addition to the first assignments made from *Ec*CsgA$_{21-151}$ (Alexander Eberth, unpublished), further 17 cross-peaks were mapped sequence specifically. The previous assignments of 13 cross-peaks to backbone amides were verified. Three repetitive units of the amyloid core were completely assigned. Thus, the backbone-resonance assignment of *Ec*CsgA was improved enabling entire H/D exchange NMR analysis and an unambiguous evaluation of the repeating units within the amyloid core.

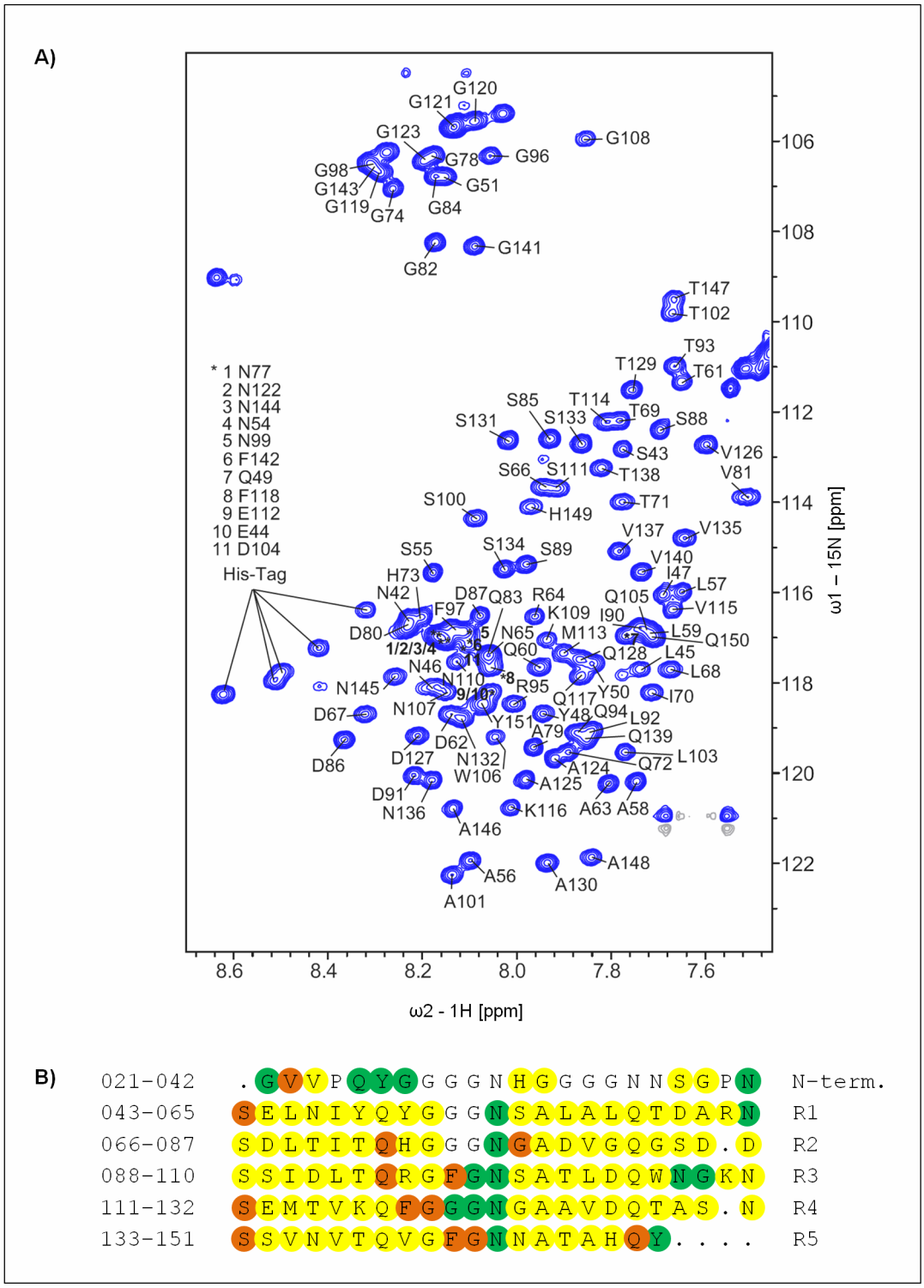

Figure 4-13: Backbone NH-assignments of *Ec*CsgA. (A) ctHSQC spectrum of uniformly ^{15}N labelled $CsgA_{40-151}$ (20 mg wet pellet of fibrils) in d_6-DMSO containing 0.1 % (v/v) d_1-TFA. Sequence-specific chemical shift assignments are indicated next to the corresponding resonance. **(B)** Improvement of the backbone assignments compared to first assignments. Confirmed assignments are mapped in yellow to the sequence, corrected assignments in orange and novel assignments in green.

H/D Exchange NMR Analysis of *Ec*CsgA$_{40\text{-}151}$ Fibrils

Since the amyloid core of *Ec*CsgA comprises the five repeating units (Collinson *et al.*, 1999) and [^{15}N,^{1}H] correlation spectra of *Ec*CsgA$_{40\text{-}151}$ showed less overlap with resonances of the highly flexible N22 tag, the truncated *Ec*CsgA variant *Ec*CsgA$_{40\text{-}151}$ was initially chosen for H/D exchange NMR analysis. ^{15}N labelled fibrils were incubated in D_2O buffer up to 20 weeks (0 h, 15 min, 1 h, 24 h, 72 h, 1 w, 4 w, 20 w) and after four weeks incubation, 45 % of the assigned resonances were reduced (**Figure 4-14**).

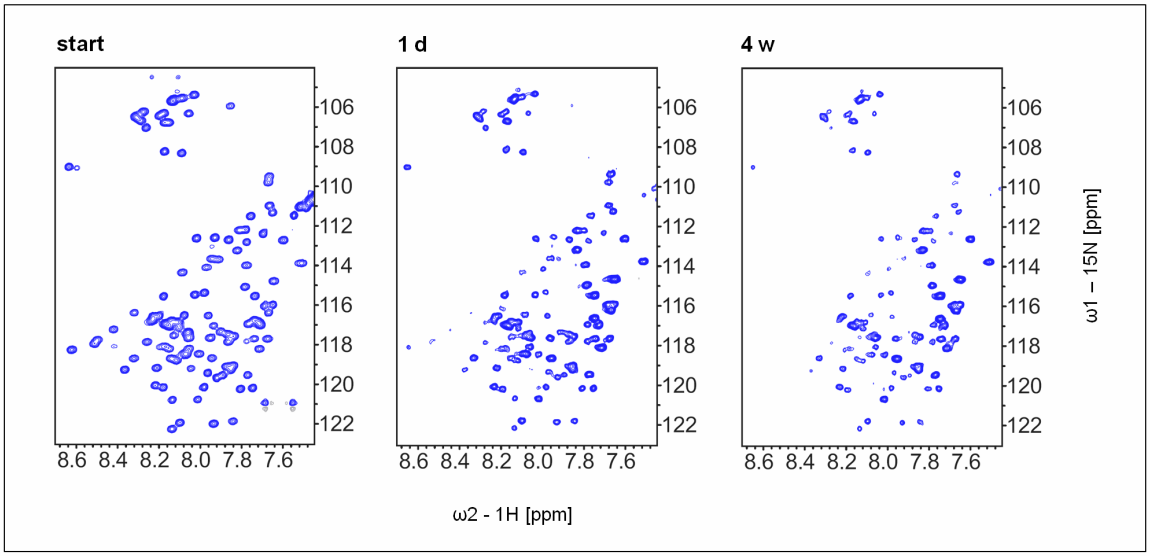

Figure 4-14: H/D exchange of *Ec*CsgA$_{40\text{-}151}$. ctHSQC spectra of uniformly ^{15}N labelled *Ec*CsgA$_{40\text{-}151}$ (~20 mg wet pellet of fibrils) in d_6-DMSO containing 0.1 % (v/v) d_1-TFA. The spectra correspond to fully protonated (start) and partially hydrogen exchanged (1 d, 4 w) *Ec*CsgA$_{40\text{-}151}$ fibrils.

Due to fast intrinsic H/D exchange in DMSO (**Figure 4-15 B**), A79 and A124 were excluded from further analysis. All other absent resonances were evaluated and the corresponding exchange rates were calculated. Nearly all residues displayed a monoexponential decay suggesting a well defined and homogeneous structure of the fibrils (**Figure 4-15 A**). Biphasic behaviour with a first fast exchanging and a second slow exchanging population was observed for the backbone amides of D67 and L68 (**Figure 4-15 C**), which are located in the second repeat, as well as A125 in the fourth and S134, G141 and A146 in the fifth repeat. The corresponding resonances are well separated and do not overlap with resonances that correspond to fast exchanging backbone amides illustrated for D67 and L68 in **Figure 4-15 D**. Furthermore, a heterogeneous behaviour of the intramolecular conserved asparagines N54, N77, N99, N122 and N144 cannot be excluded (**Figure 4-15 C**). The resonances of the backbone amides N54, N77, N122 and N144 displayed a strong overlapping. For that reason, they were analysed as one unit.

Moreover, this unit overlapped with the resonance of the backbone amide F97, which displayed fast exchange with the solvent. The resonance of N77 was separated from the other named peaks, but also overlapped with the resonance of a rapidly exchanging backbone amide, F142. Potentially, the exchange rates of the asparagines were masked by the resonance overlapping shown in **Figure 4-15 D**.

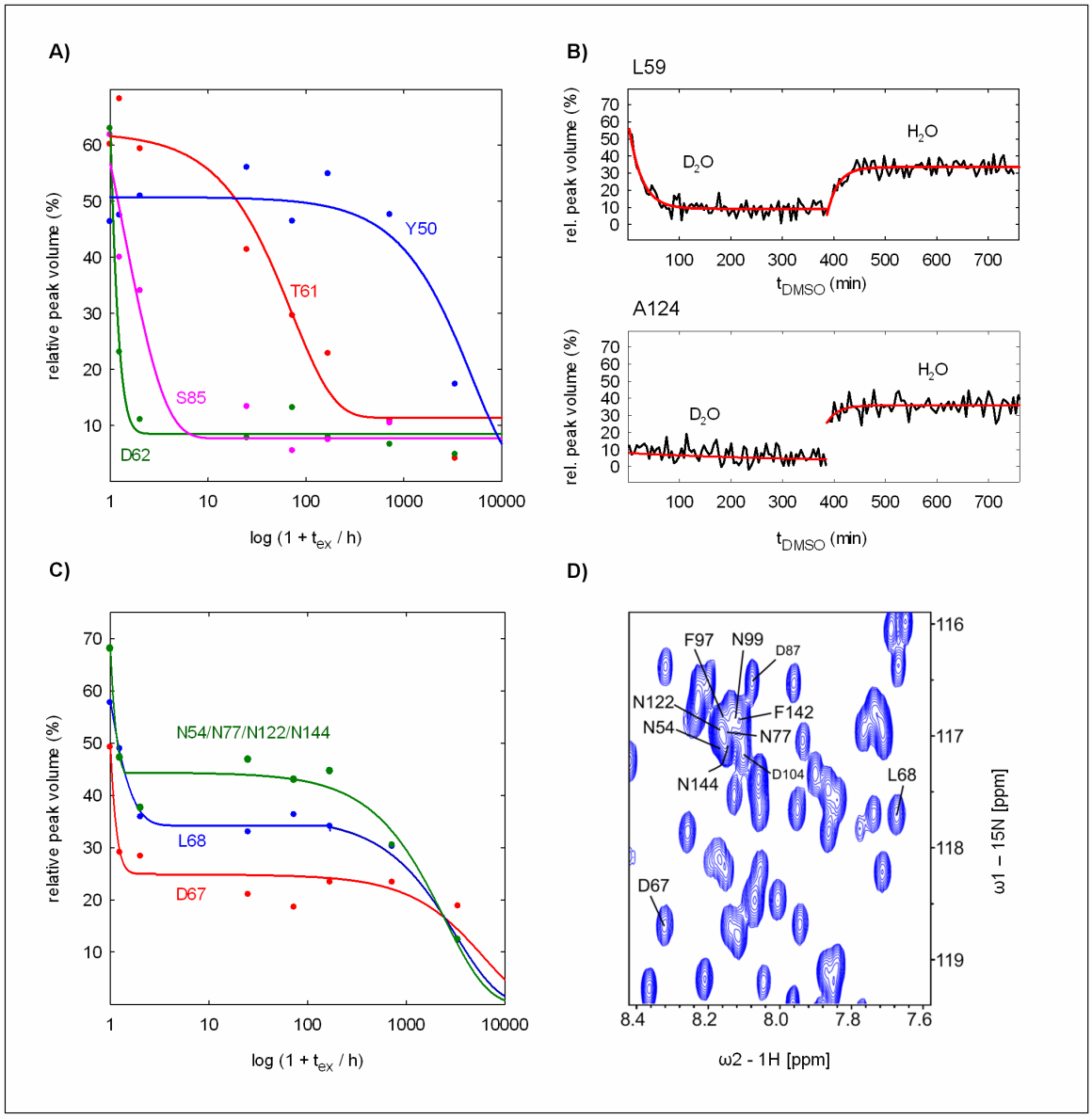

Figure 4-15: Determination of exchange rates of *Ec*CsgA$_{40\text{-}151}$. (A) Monoexponential fit of peak volumes at t_{DMSO} = 0 min to t_{ex} corresponding to selected backbone amides. **(B)** Intrinsic H/D exchange in DMSO of L59 and A124 backbone amides. The evolution of peak volumes after monomerising the *Ec*CsgA$_{40\text{-}151}$ fibrils in DMSO/TFA and final addition of H_2O is shown. Red curves show the monoexponential fits of the raw data. **(C)** Biphasic exchange behaviour of selected backbone amides. The sum of two monoexponential equations is shown. **(D)** ctHSQC spectrum of uniformly ^{15}N labelled *Ec*CsgA$_{40\text{-}151}$ (~20 mg wet pellet of fibrils) in d_6-DMSO containing 0.1 % (v/v) d_1-TFA. The section shows the backbone amides of the intramolecularly conserved asparagines, as well as D67 and L68 that display biphasic behaviour in H/D exchange.

The resulting exchange rates are shown in **Figure 4-16**. In case of biphasic behaviour, nearly all appropriate residues featured a major slow and a minor fast exchanging population (**Figure App-6-1**). The H/D exchange NMR analysis identified two segments per repeat that displayed slow exchange rates of 10^{-5} h $^{-1}$ to 10^{-1} h^{-1}. Similar exchange rates have been found for the amyloid core of *Pa*HET-$s_{218-289}$ (Ritter *et al.*, 2005) and $A\beta_{1-42}$ fibrils (Luhrs *et al.*, 2005) indicating the high stability and compactness of the fibrils. These segments, which comprise the backbone amides of the residues S43-Y50, S55-T61, D67-H73, G78-G84, S89-R95, S100-W106, E112-F118, G123-Q128, S134-V140 and N145-Y151, were assumed to be involved in hydrogen bonds that protect the corresponding backbone amides against hydrogen exchange. As the biophysical characterisation of *Ec*$CsgA_{40-151}$ (**chapter 4.3.4**) confirmed a β-sheet rich amyloid structure, the protected segments were assigned to β-strands that consist of seven to eight residues. The conserved serines (S43, S66, S88, S111 and S133) exhibited different exchange behaviour and the conserved asparagines (N54, N77, N88, N122 and N144) seemed to be masked. Thus, their β-strand incorporation remains elusive. In contrast to the protected backbones, the backbone amides of N42, G51, D62-S66, G74, S85-S88, G96-G98, N107-N110, G119-G121, T129-S133 and G141-G143 exchanged rapidly ($k \geq 1.5$ h^{-1}). Since these backbone amides are only weak or not protected against H/D exchange with the solvent, they seem to be conformationally disordered. In consequence, the corresponding residues were identified to form turn and loop regions that connect the distinct β-strands.

Thus, the H/D exchange analysis of *Ec*$CsgA_{40-151}$ identified every repeating unit to form two β-strands that are connected by a short β-turn. The β-strands of adjacent repeats are connected by flexible loops comprising four to five residues. The protected segments 1, 3, 5, 7, 9 and 2, 4, 6, 8, 10 in each case are separated by 22-23 residues indicating a repetitive structural pattern.

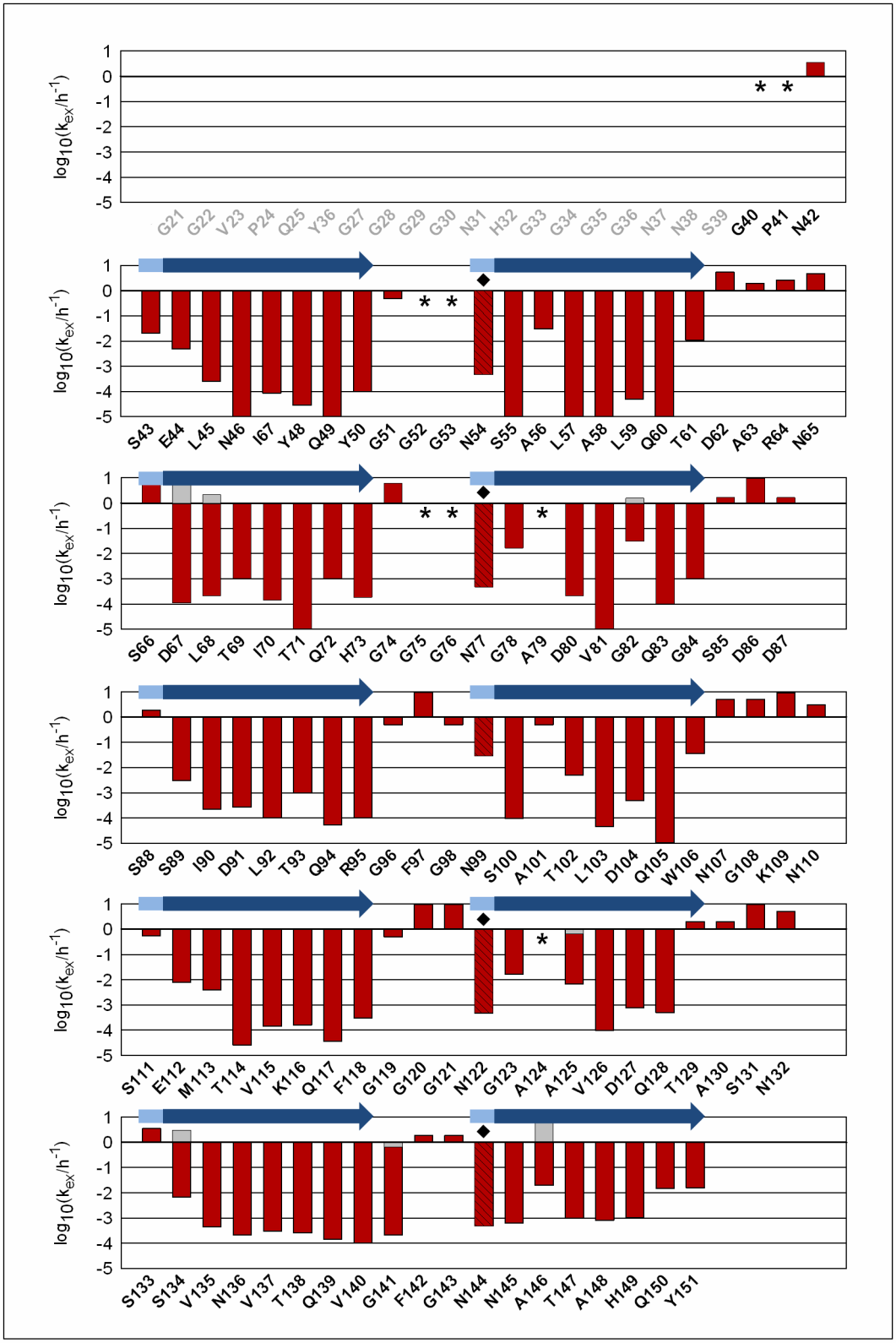

Figure 4-16: H/D exchange NMR data of *Ec*CsgA$_{40\text{-}151}$ and secondary structure prediction. Red and grey bars indicate exchange rates of *Ec*CsgA$_{40\text{-}151}$ k_{ex} (h^{-1}). Grey bars indicate exchange rates of the fast first phase in case of biphasic behaviour. Patterned bars present the exchange rates of the intramolecularly conserved asparagines. Residues marked with a rectangle were analysed as one unit. Residues marked with an asterisk were excluded from the analysis due to fast intrinsic exchange in DMSO (A79, A124) and missing assignments. Blue arrows indicate β-strands. The β-strand incorporation of the intramolecularly conserved serines and asparagines remained unclear shown in light-blue.

H/D Exchange NMR Analysis of *Ec*CsgA$_{21-151}$ Fibrils

Since *in vivo* mature *Ec*CsgA comprises the residues 21 to 151, wt *Ec*CsgA$_{21-151}$ fibrils were subjected to H/D exchange analysis. The mature full-length protein was probed in the same context to validate the results of the H/D exchange NMR experiments of *Ec*CsgA$_{40-151}$ fibrils. Do the fibrils feature the same structural pattern and same homogeneity? ^{15}N labelled *Ec*CsgA$_{21-151}$ fibrils were processed in the same manner as *Ec*CsgA$_{40-151}$ and H/D exchange was followed up to one week (0 h, 1 h, 1 d and 1 w) illustrated in **Figure 4-17**.

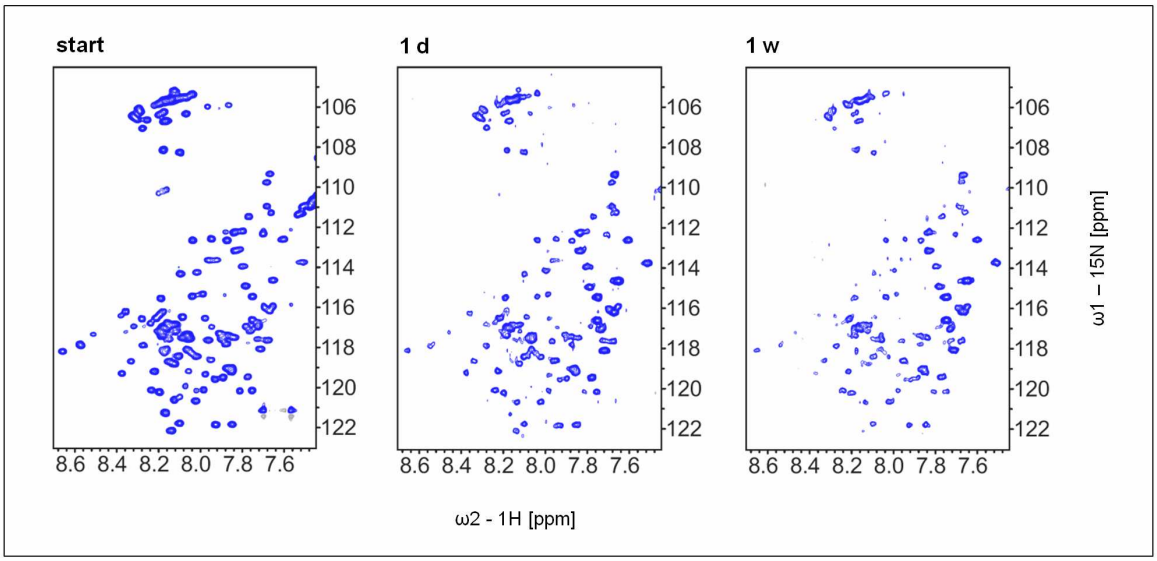

Figure 4-17: H/D exchange of *Ec*CsgA$_{21-151}$. ctHSQC spectra of uniformly ^{15}N labelled *Ec*CsgA$_{21-151}$ (~20 mg wet pellet of fibrils) in d_6-DMSO containing 0.1 % (v/v) d_1-TFA. The spectra correspond to fully protonated (start) and partially hydrogen exchanged (1 d, 1 w) *Ec*CsgA$_{21-151}$ fibrils.

In addition to A79 and A124, the resonances of N46 and A58 were excluded from further analysis due to fast H/D exchange in DMSO. Thus, H/D exchange rates were determined for all assigned residues except N46, A58, A79 and A124. Most of the backbone amides featured a monophasic H/D exchange reaction similar to *Ec*CsgA$_{40-151}$. In comparison to the truncated variant, D67, L68, A125 and the asparagines N54, N77, N122 and N144 demonstrated the same biphasic reaction as in *Ec*CsgA$_{40-151}$ fibrils. In contrast to the analysis of *Ec*CsgA$_{40-151}$, N99, S134, G141 and A146 revealed less biphasic H/D exchange reaction shown for N99 and A146 in **Figure 4-18 A**. Furthermore, biphasic behaviour was observed for the resonances of the backbone amides I70, T71 and G78 that are located in the second repeat, I90 in the third repeat, as well as M113, T114, G123 and D127, which are located in the fourth repeat (**Figure 4-18 B**). The corresponding resonances are mostly well separated. The resonances of I90, G78 and G123 demonstrated overlap with other

resonances illustrated in **Figure 4-18 C/D**. I90 overlapped with resonances of slow exchanging backbone amides, and G78 and G123 overlapped with each other. Since the spectra of *Ec*CsgA$_{21\text{-}151}$ featured resonances of the glycine rich N22 tag, overlap with a non assigned flexible glycine cannot be excluded. G78, G123 and D127 are amino acids, which have been assigned to "gatekeeper" residues before (Wang *et al.*, 2010). Nearly all biphasic exchanging residues revealed a major slow and a minor fast exchanging population (**Figure App-6-1**).

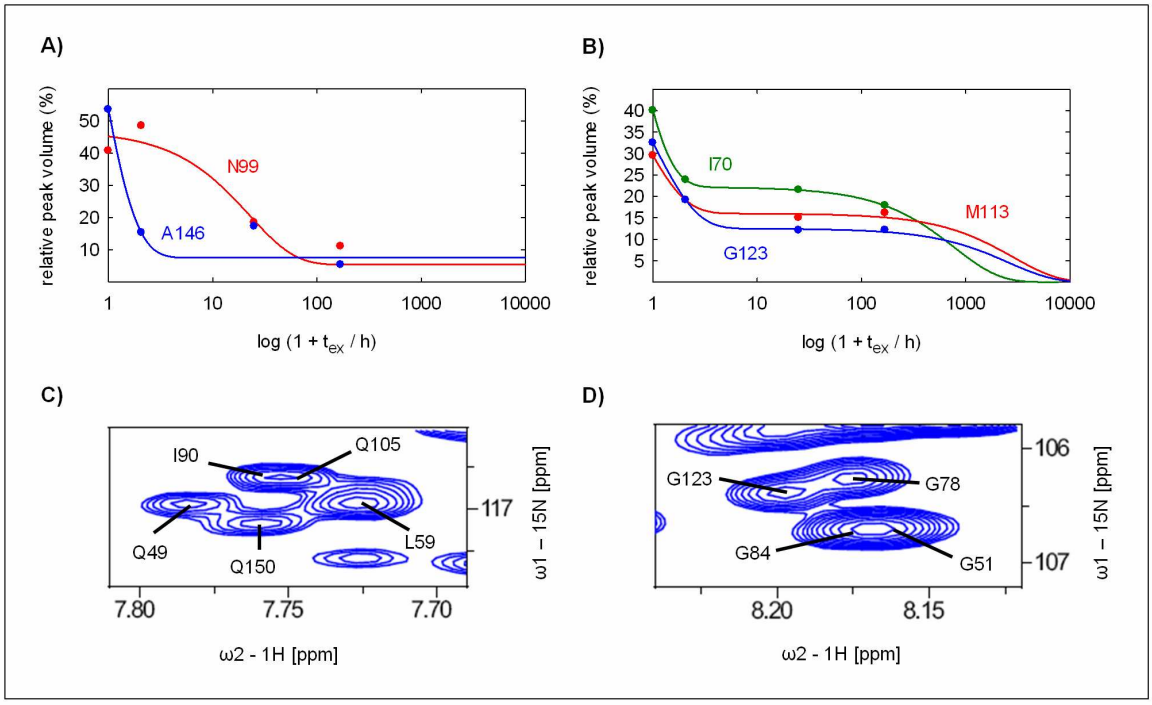

Figure 4-18: Determination of exchange rates of *Ec*CsgA$_{21\text{-}151}$. (A) Monoexponential fit of peak volumes at t_{DMSO} = 0 min to t_{ex} corresponding to selected backbone amides. **(B)** Biphasic exchange behaviour of selected backbone amides. The sum of two monoexponential equations is shown. **(C)** and **(D)** ctHSQC spectra of uniformly ^{15}N labelled *Ec*CsgA$_{21\text{-}151}$ (~20 mg wet pellet of fibrils) in d6-DMSO containing 0.1 % (v/v) d_1-TFA. The respective section shows the backbone amides of I90, G78, G123 and overlapping resonances.

Figure 4-19 presents the observed exchange rates for *Ec*CsgA$_{21\text{-}151}$ fibrils in comparison to the exchange rates of *Ec*CsgA$_{40\text{-}151}$. The slow exchange rates of *Ec*CsgA$_{21\text{-}151}$ were masked due to the short H/D exchange period (≤ 1 w). In most cases, the short exchange period prohibited the examination of exchange rates $k \leq 10^{-2}$ h^{-1}. Therefore, the exchange rates are in a range of 10^{-2} h^{-1} to 10 h^{-1}. Supplementary, **Figure App-6-2** illustrates the results for *Ec*CsgA$_{21\text{-}151}$ solely.

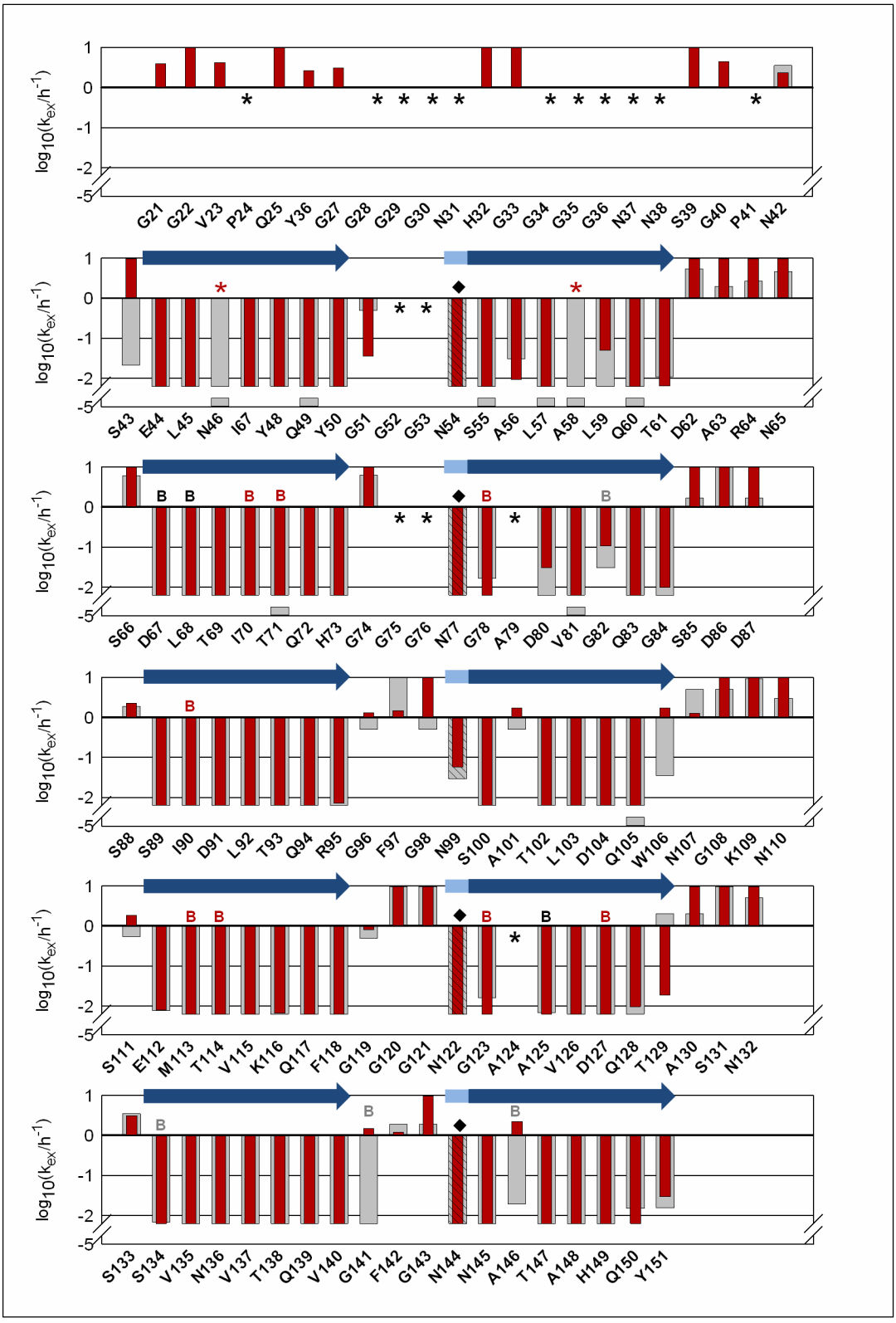

Figure 4-19: H/D exchange NMR data of *Ec*CsgA$_{21\text{-}151}$ and secondary structure prediction. Red bars indicate exchange rates of *Ec*CsgA$_{21\text{-}151}$ k_{ex} (h^{-1}), grey bars indicate exchange rates of *Ec*CsgA$_{40\text{-}151}$ k_{ex} (h^{-1}). Patterned bars present the exchange rates of the intramolecularly conserved asparagines. Residues marked with a rectangle were analysed as one unit. Residues marked with an asterisk were excluded from the analysis due to fast intrinsic exchange in DMSO (N46, A58, A79, and A124) and missing assignments. Blue arrows indicate β-strands. The β-strand incorporation of the intramolecularly conserved asparagines remained unclear shown in light-blue.

The H/D exchange analysis of *Ec*CsgA$_{21-151}$ fibrils revealed ten segments that displayed slow exchange rates. Similar to *Ec*CsgA$_{40-151}$ fibrils, the backbone amides of E44-Y50, S55-T61, D67-H73, G78-G84, S88-R95, N99-W105, S111-F118, G123-Q128, S134-V140 and N145-Y151 were protected against H/D exchange and thus assigned to β-strands. The assigned backbone amides of the N22 tag, S43, G51, D62-S66, G74, S85-S87, G96-G98, W106-N110, G119-G121, A130-S133 and G141-G143 exchanged rapidly. Thus, they are highly flexible.

In contrast to the H/D exchange analysis of *Ec*CsgA$_{40-151}$, all intramolecular conserved serines rapidly exchanged and were thus excluded from the β-strands. The β-strand incorporation of the intramolecular conserved asparagines N54, N77, N99, N122 and N144 remained still elusive due to the strong overlapping (**Figure 4-15 D**). However, the backbone-resonance of N99 disappeared in a less biphasic manner as before and the slow monoexponential exchange rate indicated the affiliation to the corresponding β-strand. In consideration of the repetitive structural pattern and the presumed formation of a stacking asparagine ladder, the corresponding asparagines should behave uniformly. In that case, the biphasic intensity decrease was induced by overlap with a fast exchanging residue and not by different exchange behaviour of the asparagines. Thus, a β-strand conformation of all appropriate asparagines is conceivable.

The backbone amides of the intramolecular conserved alanines A101 and A146 exhibited a fast exchange in a protected region indicating a structural disturbance at these positions. *Ec*CsgA$_{40-151}$ revealed also less protection or biphasic H/D exchange behaviour at these positions. Unfortunately, exchange data are not available for A79 and A124 due to fast exchange in DMSO. Even A56 is protected against H/D exchange, a structural disturbance like a kink or a β-arc at the positions of the conserved alanines is feasible. The precedent H/D exchange analysis of *Fg*HET-s$_{218-289}$ fibrils also did not reveal fast exchange at all β-arc positions (**chapter 4.2.3**, **Figure 4-7**).

In summary, the H/D exchange analysis of *Ec*CsgA$_{21-151}$ fibrils identified a repetitive structural pattern with a two β-strands per repeating unit congruent to the results of *Ec*CsgA$_{40-151}$. In comparison to *Ec*CsgA$_{40-151}$, the analysis revealed a biphasic exchange behaviour at similar positions in most instances. However, the exchange behaviour was defined and supported a specific fibrillar structure of both variants. Thus, the analysis is fully congruent to the analysis of *Ec*CsgA$_{40-151}$ fibrils. Furthermore, the exchange data indicated that the N22 tag is highly flexible. Hence, the exchange pattern confirms the

domain character of mature *Ec*CsgA (21-151), which consists of a flexible protease-susceptible and a well structured protease-resistant domain (Collinson *et al.*, 1999).

H/D Exchange Analysis of Heteronucleated Fibrils

In vivo the fibrillisation of CsgA is nucleator dependent. CsgB is indispensible for the assembly of monomeric CsgA into appropriate amyloid fibrils (Hammar *et al.*, 1996; Hammer *et al.*, 2007).

In order to investigate the effect of the heteronucleation to the *Ec*CsgA structure and to investigate a detailed molecular model of the nucleator dependent assembly, ^{15}N labelled *Ec*CsgA$_{40\text{-}151}$ was seeded by 10 % (w/w) unlabelled *Ec*CsgB fibrils and the newly formed *Ec*CsgA$_{40\text{-}151}$ fibrils were analysed by H/D exchange NMR. The H/D exchange reaction was quenched after 15 min, 1 h, 1 d, 1 w and 4 w. After incubation in D_2O buffer for four weeks, already 60 % of the assigned resonances of *Ec*CsgA$_{40\text{-}151}$ were significantly reduced or absent illustrated in **Figure 4-20**.

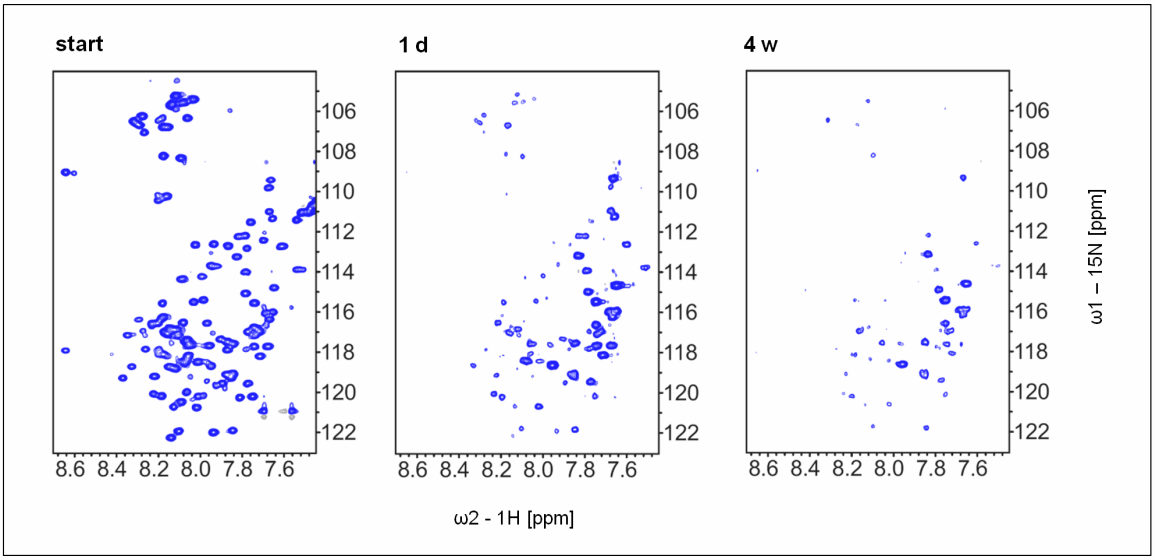

Figure 4-20: H/D exchange of *Ec*CsgA$_{40\text{-}151}$ after seeding with *Ec*CsgB. ctHSQC spectra of uniformly ^{15}N labelled *Ec*CsgA$_{40\text{-}151}$ (~25 mg wet pellet of fibrils) after seeding with 10 % (v/v) *Ec*CsgB in d_6-DMSO containing 0.1 % (v/v) d_1-TFA. The spectra correspond to fully protonated (start) and partially hydrogen exchanged (1 d, 4 w) *Ec*CsgA$_{40\text{-}151}$ fibrils.

In contrast to the precedent analysis of *Ec*CsgA$_{40\text{-}151}$ and *Ec*CsgA$_{21\text{-}151}$, remarkable 30 % of the residues revealed an H/D exchange in a biphasic manner indicating heterogeneity. The relative exchange behaviour as well as the relative population of *Ec*CsgA$_{40\text{-}151}$ fibrils after seeding with 10 % (w/w) *Ec*CsgB fibrils is shown in **Figure 4-21**. The detailed exchange rates are shown in the Appendix in **Figure App-6-3**.

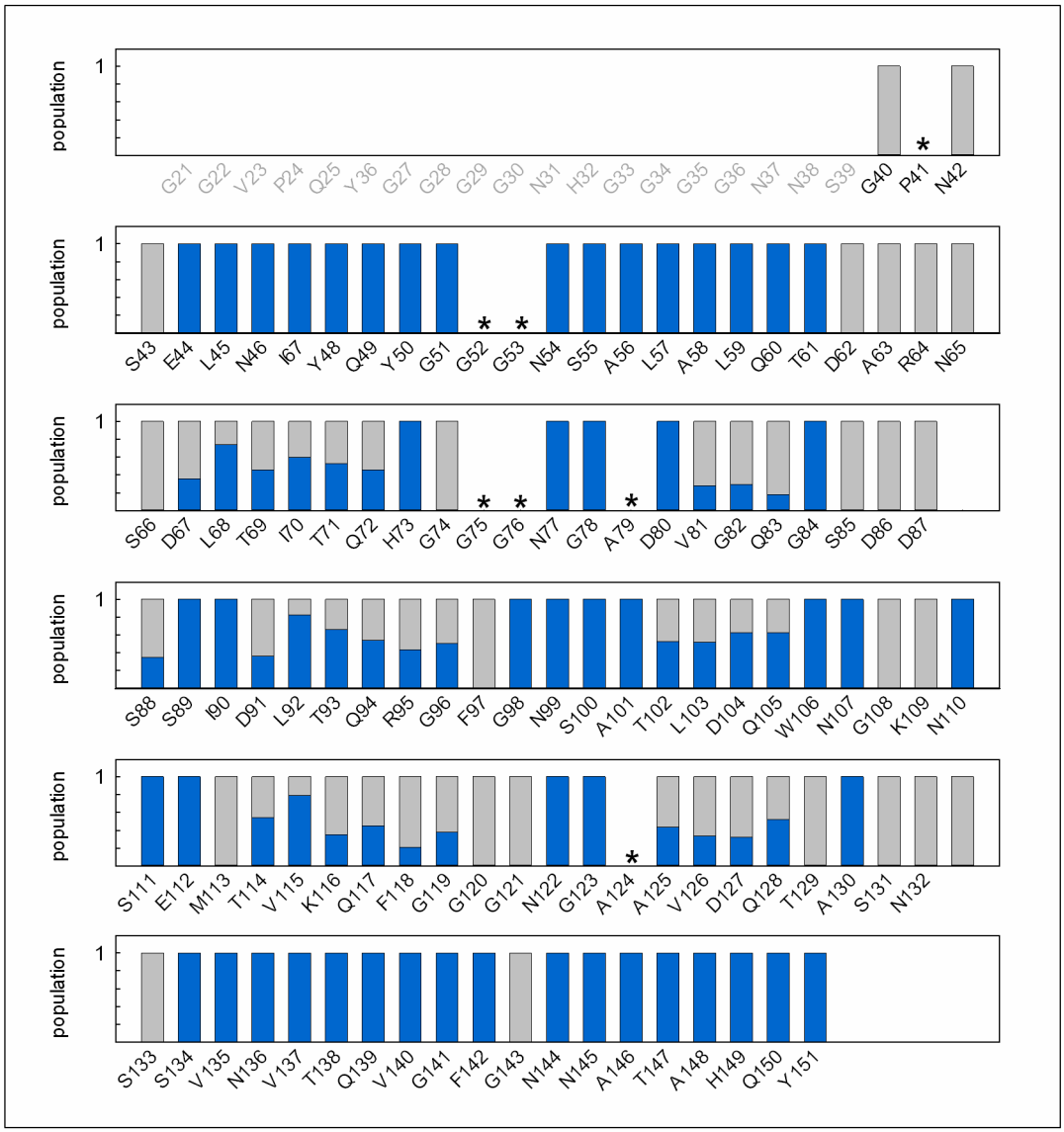

Figure 4-21: Relative exchange behaviour and population of *Ec*CsgA$_{40\text{-}151}$ after seeding with *Ec*CsgB fibrils. Fast exchange is shown in grey, slow exchange in blue. Residues marked with an asterisk were excluded from the analysis due to fast exchange in DMSO (A79, A124) and missing assignments.

The backbone amides of the N-terminal residues, as well as the first and fifth repeat displayed a homogeneous protection pattern, which is congruent to the analysis of *Ec*CsgA$_{40\text{-}151}$ and *Ec*CsgA$_{21\text{-}151}$. Both, first and fifth repeat revealed two protected segments per repeat. But in contrast to homonucleated *Ec*CsgA$_{40\text{-}151}$ fibrils, most slow exchange rates are in a range of 10^{-3} h^{-1} to 10 h^{-1} indicating less stability. The exchange rates of the central repeats (repeat 2, 3 and 4) exhibited a less regular protection pattern. Nearly 50 % of the assigned residues in this segment showed biphasic H/D exchange behaviour with a fast and a slow exchanging population. Less of these backbone amides revealed a major population, which exchanged slowly. In most instances, the major population rapidly exchanged. The exchange behaviour of backbone amides, which

displayed an almost exclusively monoexponential exchange, was congruent to former analysis of *Ec*CsgA$_{40-151}$ and *Ec*CsgA$_{21-151}$ except, G98, N107, N110 and A130. These residues displayed an unexpected slow exchange within weak protected segments. Since half of the backbone amides in the central repeats exchanged in a biphasic manner without any regularity and even monophasic amides revealed an unexpected exchange rate, the structure of the central repeats is very heterogeneous and less regular.

Recapitulatory, the H/D exchange analysis of *Ec*CsgA$_{40-151}$ fibrils, which were nucleated by 10 % (v/v) *Ec*CsgB fibrils, revealed a homogeneous exchange behaviour for the first and fifth repeat and a heterogeneous and less regular structure of the three central repeats. Thus, there are distinct differences in comparison to *Ec*CsgA fibrils, which were nucleated by preformed *Ec*CsgA$_{21-151}$ fibrils.

4.3.7 Characterisation of the β-Sheet Architecture

Owing to their natural properties, β-sheets always feature a specific architecture. Every amyloid core exhibits solvent accessible and non accessible side-chains. The arrangement of the β-strands enables characteristic inter- and intramolecular contacts. In order to analyse the topology of β-strands as well as these inter- and intramolecular contacts within *Ec*CsgA$_{21-151}$ fibrils, cysteine variants of *Ec*CsgA$_{21-151}$ were subjected to thiol-reactive probes.

Initial Labelling Experiments

In labelling experiments appropriate cysteine variants (**Table 3-7**) were incubated in presence of the thiol-reactive probes iodaceteic acid (IAA), iodacetamide (IAM), N-(1-pyrene)maleimide (PM) and AlexaFluor488 C_5 maleimide (AF488) under numerous conditions as described in **chapter 3.6.3**. Thus, the fibrillisation- and labelling conditions of *Ec*CsgA$_{21-151}$ cysteine variants were optimised. Finally, the variants were co-aggregated with 10 % (w/w) wt *Ec*CsgA$_{21-151}$ and seeded with preformed *Ec*CsgA$_{21-151}$ fibrils. Reducing fibrillisation and labelling conditions emerged as indispensable (**Figure 4-22**). The double variant *Ec*CsgA$_{21-151}$ E44C/N46C, which was expected to feature two solvent accessible cysteines close to each other, was chosen for initial PM-labelling studies. PM-labelling can lead to the formation of excimers, which are detected by a strong red

shift in fluorescence (Lehrer, 1997). This typical red shift was only observed after fibrillisation and labelling under reductive conditions (**Figure 4-22**).

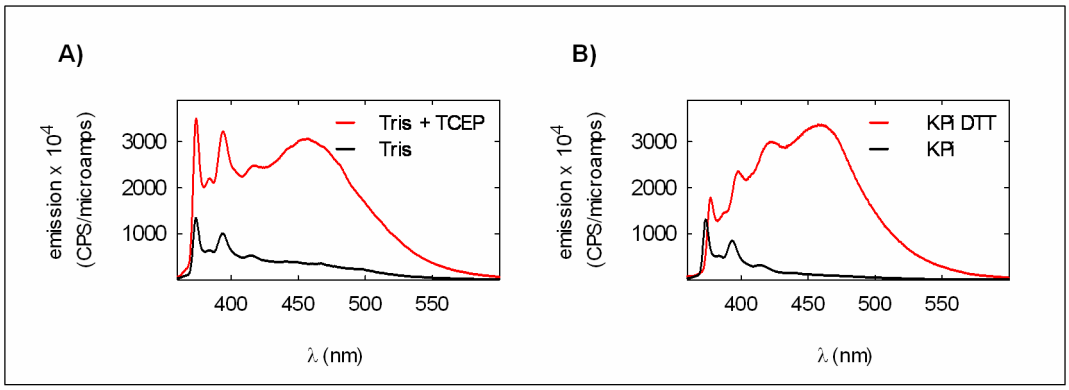

Figure 4-22: Pyrene excimer formation of *Ec*CsgA$_{21-151}$ E44C/N46C fibrils depending on reductive fibrillisation conditions. Approximately 15 µM fibrils after PM labelling were applied to fluorescence spectroscopy. Covalent bound pyrene was excited at λ_{ex} = 342 nm. Protein samples were fibrillised in presence and absence of 5 mM TCEP **(A)** and 5 mM DTT **(B)**.

The reductive conditions avoided disulphide formation and facilitated free accessible thiol groups that bound PM. The proximity of the cysteines enabled pyrene excimer formation and thus led to a characteristic red shift fluorescence at ~ 455 nm (Lehrer, 1997). In preparation of further labelling experiments, cysteine variants were always fibrillised in the presence of 5 mM DTT according to the conditions mentioned in **chapter 3.6.3**.

After the labelling reaction proteins were applied to mass-spectrometry to confirm the specific binding of the thiol-reactive probes. Since the thiol-labelling is covalent, the reaction results in a mass shift. **Figure 4-23** shows mass spectra of *Ec*CsgA$_{21-151}$ V23C after labelling reactions with IAM, PM and AF488. The residue 23 is located in the flexible N terminus and is thus expected to be solvent accessible. Considering N-formylation (Δm = 28 Da per N-formylation) and a probably missing start methionine (Δm = 149.2 Da), an appropriate mass shift was only detected after labelling with PM (Δm = 297 Da) and AF488 (Δm = 643 Da). Owing to the missing mass shift, IAM was not binding at all neither to fibrillar nor unfolded *Ec*CsgA$_{21-151}$ cysteine variants. A similar behaviour was observed for IAA (data not shown). Thus, IAM and IAA were excluded from further analysis.

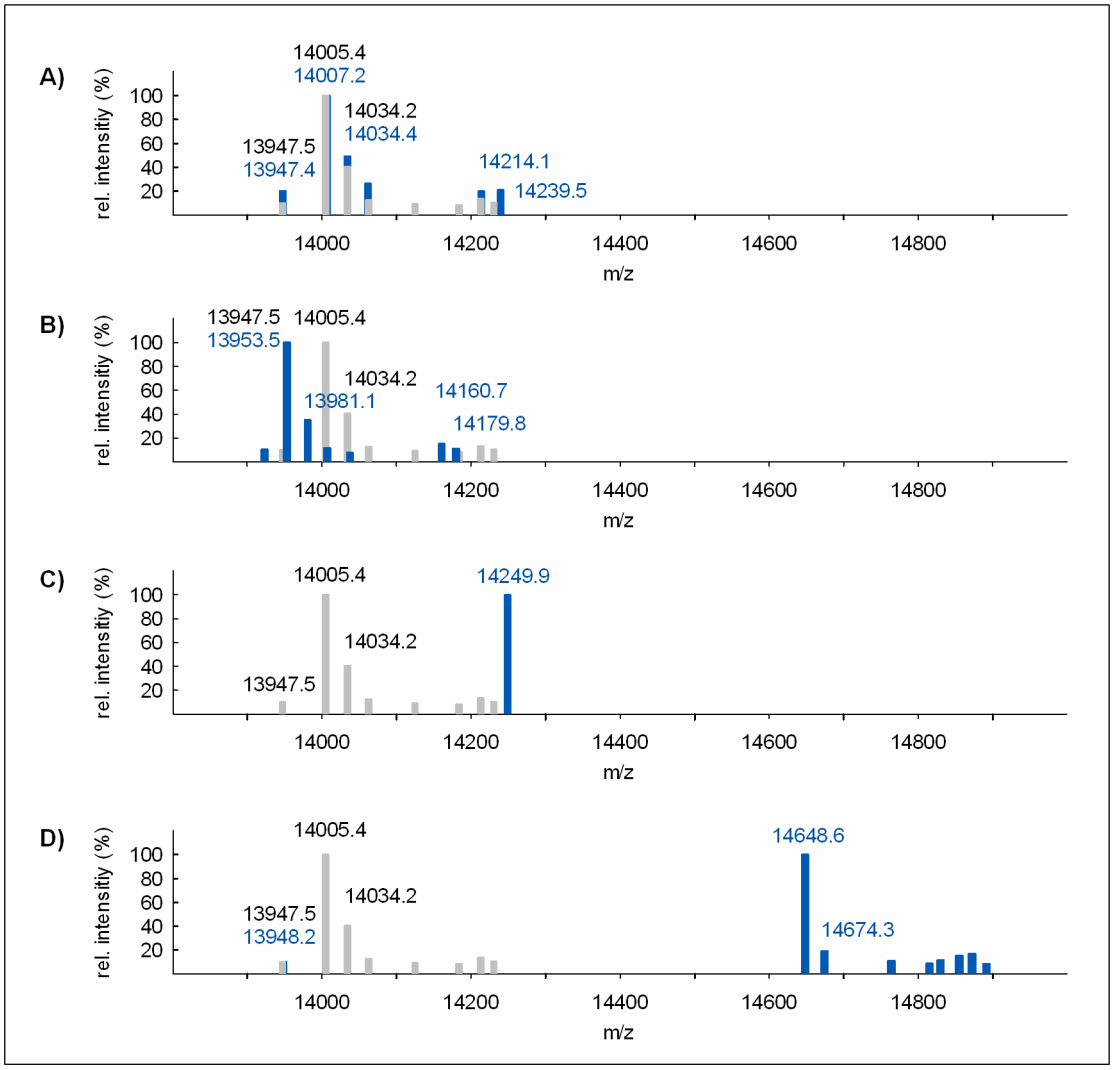

Figure 4-23: MALDI spectra after thiol-labelling of *Ec*CsgA$_{21\text{-}151}$ V23C. Approximately 50 µg *Ec*CsgA$_{21\text{-}151}$ V23C were applied to mass analysis after thiol-labelling. Owing to thiol-labelling, protein samples were incubated in presence of IAM under native conditions **(A)** and denaturing conditions **(B)**, PM **(C)** and AF488 **(D)** under native conditions. Mass of *Ec*CsgA$_{21\text{-}151}$ V23C prior labelling is shown in grey, mass of probably labelled protein is shown in blue. Calculated masses are 14049.8 Da for unlabelled protein, 14234.8 Da for covalent IAM labelling, 14347.1 Da for covalent PM labelling and 14692.8 Da for covalent AF488 labelling. The observed masses were masked due to N-formylation and a probably absent start methionine.

Side-Chain Solvent Accessibility

In order to characterise the β-strand topology of *Ec*CsgA fibrils and to identify suitable cysteine variants for the analysis of intra- and intermolecular contacts within *Ec*CsgA fibrils, appropriate fibrils of *Ec*CsgA$_{21\text{-}151}$ cysteine variants were applied to thiol-labelling with N-(1-pyrene)maleimide PM. Fibrils were processed and analysed by mass spectrometry as described in **chapter 3.6.3**.

Considering N-formylation and a probably missing start methionine, MALDI spectra revealed an appropriate mass shift for numerous variants (**Figure App-6-4**) indicating a

specific labelling and thus side-chain solvent accessibility. Illustrated in **Figure 4-24**, 14 of the 17 cysteine variants, which carried a single cysteine residue in the amyloid core region, were labelled with PM.

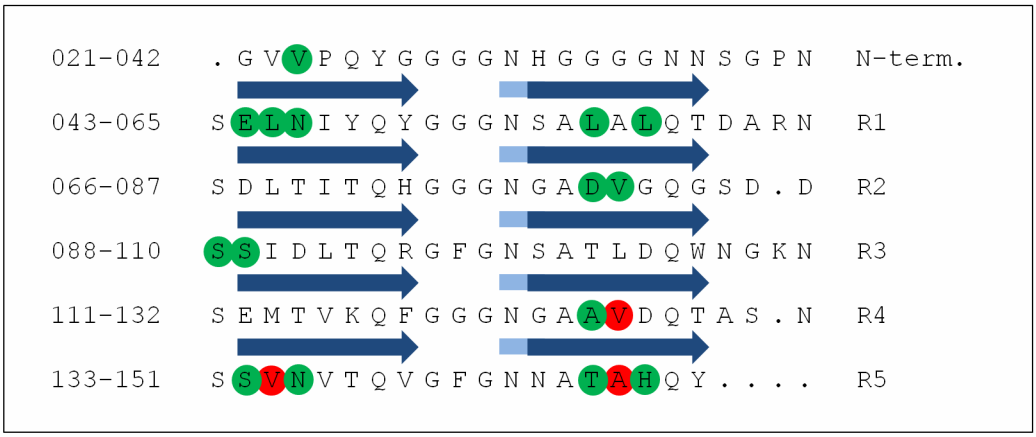

Figure 4-24: Solvent accessibility of *Ec*CsgA$_{21-151}$ single cysteine variants. The relative labelling efficiency is shown for appropriate variants. Solvent accessible side-chains are highlighted in green, not accessible residues in red. Blue arrows indicate β-strands, which were localised by H/D exchange NMR in this study.

The high cross-linking rates of E44C, N46C, L57C, L59C, D80C, S89C, A125C, S134C, N136C, T147C and H149C, as well as the low rates of V81C, V135C and A148C are fully congruent to the repetitive structural pattern of *Ec*CsgA indicated by H/D exchange analysis and the domain character of the protein. The side-chain solvent accessibility of the analysed residues is fully congruent to the presumption that amyloids feature solvent accessible and non accessible side-chains in a defined manner. Furthermore, the results indicated that all β-sheets feature one solvent accessible side. Considering this fact, the β-strand topology and the repetitive character of *Ec*CsgA, L45C and V81C behaved contrariwise. The corresponding cysteine residues were accessible. Due to the natural topology of β-strands it is not possible that three adjacent side-chains feature the same orientation. Thus, the accessibility can only be explained by structural differences of the cysteine variants. The neighbouring serine residues S88 and S89 were both identified as solvent accessible. In conformity with the H/D exchange analyses, S88 is part of a loop region leading into the β-strand, which incorporates S89. Considering the geometry of turns and loops, both residues can feature solvent accessibility.

In addition to PM-labelling, side-chain solvent accessibility was also probed by AF488 labelling for selected cysteine variants. The *Ec*CsgA$_{21-151}$ variants V23C, E44C, N46C and N136C were labelled with AF488 and processed as described in **chapter 3.6.3**. Congruent to the PM labelling, MALDI spectra demonstrated the expected mass shift owing to covalent thiol modification illustrated in **Figure 4-25**.

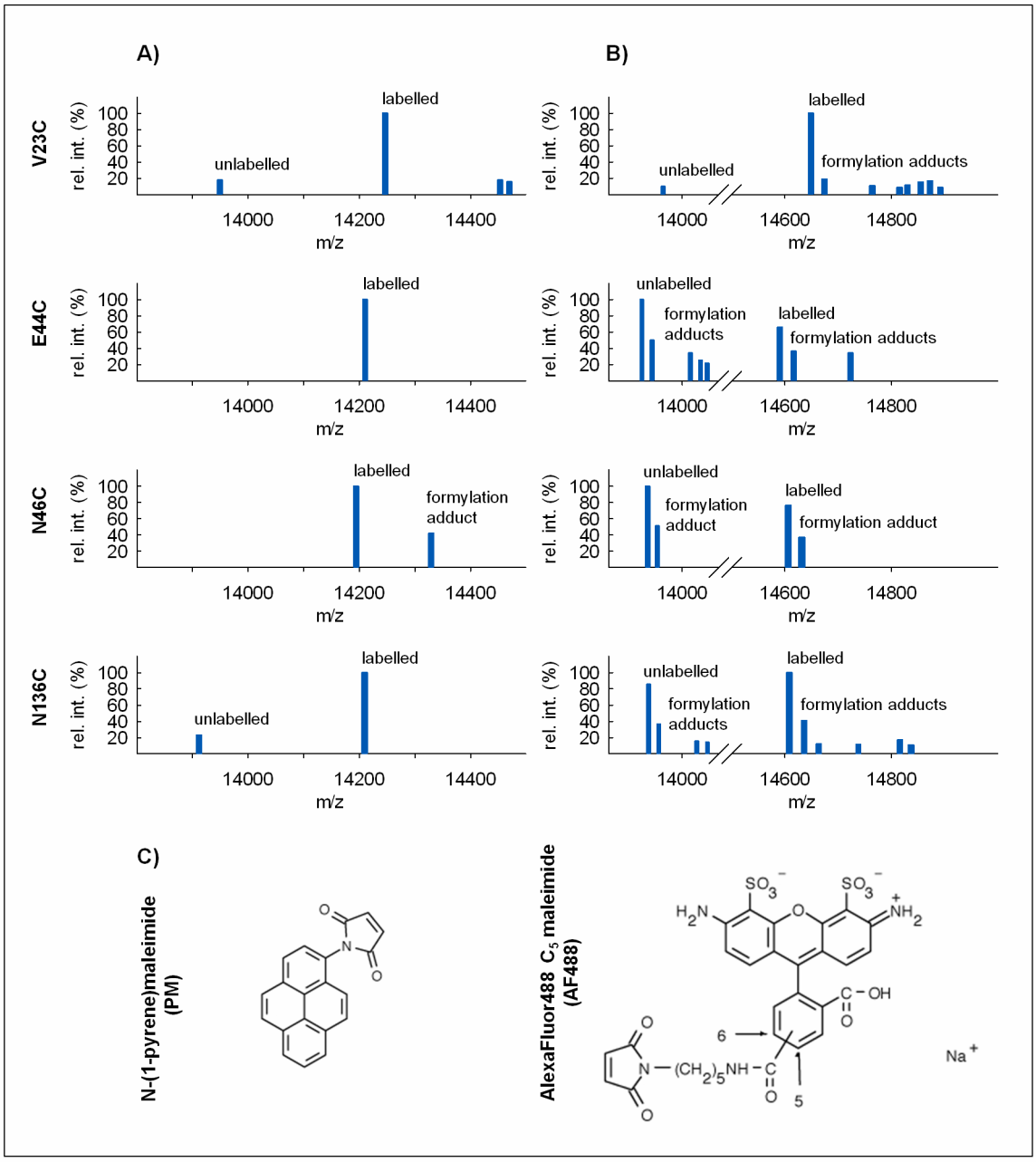

Figure 4-25: MALDI spectra after thiol-labelling of *Ec*CsgA$_{21-151}$ cysteine variants. Approximately 50 µg of *Ec*CsgA$_{21-151}$ cysteine variants were subjected to mass analysis after thiol-labelling. **(A)** MALDI spectra after PM labelling. **(B)** MALDI spectra after AF$_{488}$ labelling. **(C)** Chemical structures of PM and AF$_{488}$.

V23C, which carries the cysteine residue in the highly flexible N-terminal tail, was completely labelled with PM and AF$_{488}$. In contrast to PM labelling, after AF$_{488}$ labelling unlabelled protein was detected for the variants E44C, N46C and N136C indicating an incomplete reaction. These variants bore the cysteine residues in the amyloid domain. Despite solvent accessibility, these residues might be even less accessible for thiol-reactive probes depending on the size of the chemical compound and the arrangement of the protofilaments within the fibril. Since AF$_{488}$ and PM differ in their size, a different

labelling efficiency is conceivable. Furthermore, this fact indicates the high compactness of the amyloid fibrils of *Ec*CsgA.

4.4 Stable Oligomers – Initial Experiments on Analogues of CsgA Oligomers

Since oligomeric intermediates of the fibrillisation pathway are believed to be the toxic agent in case of disease related amyloid formation, their structure-function relationship is of high interest. In order to obtain soluble analogues of *Ec*CsgA oligomers, three general ideas were followed up. The principle was the negative design of natural soluble β-sheet proteins to avoid edge-to-edge aggregation (Richardson and Richardson, 2002). Due to the natural occurring edge strands, which exhibit free hydrogen bond donors or acceptors, β-sheets are always aggregation prone. Natural soluble β-sheet proteins are protected by different features against the interaction with any other β-strands.

The following natural strategies were chosen to inhibit CsgA aggregation:

a) Introduction of the β-breaker proline.
b) Introduction of an inward pointing charged side chain by lysine.
c) Covering the edge strands with a loop of a helical or nonrepetitive structure by fusing the well folded small proteins barnase and sumo.

4.4.1 Design and Expression of CsgA Fusion Constructs

All constructs in that approach were made from the repetitive core of *Ec*CsgA lacking the N22 tag. Following the concept of the negative design by introducing proline or lysine into predicted edge strands, additional repeating units were fused. These additional strands should act as proper edge strand, thus named caps, and were modified accordingly. Therefore and to create duplicates of the core, R5 was extended to the same length of the other repeating units. R5 was prolonged by the addition of amino acids that correspond to the loop regions of the other repeats. As further fusion partners, an inactive variant of barnase from *Bacillus amyloliquefaciens* (Schmoldt *et al.*, 2005) and sumostar were used. All corresponding gene constructs were generated as described in **chapter 3.3.4**. Encoded proteins are shown in **Figure 4-26 A**.

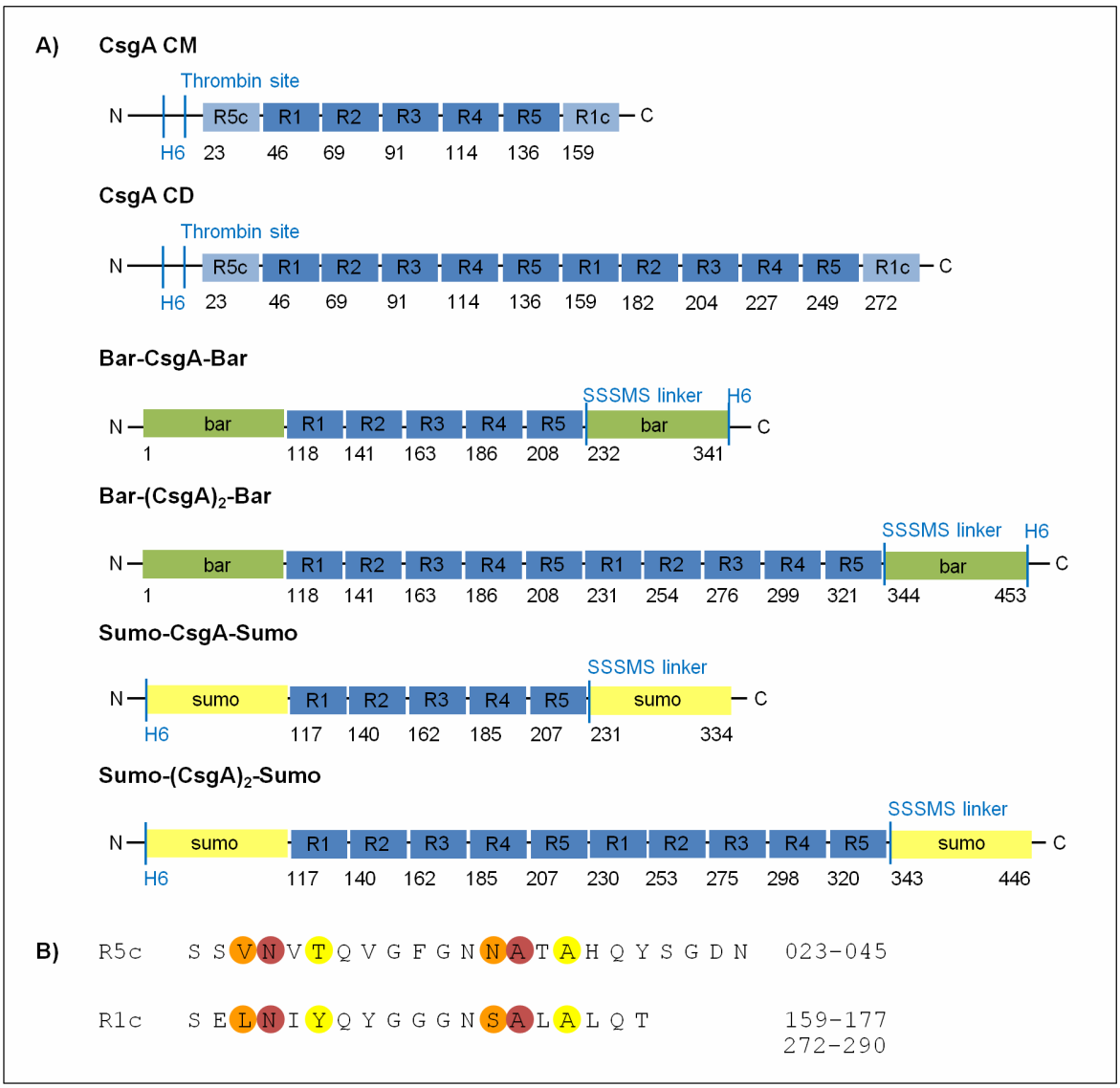

Figure 4-26: Design of CsgA fusion proteins. (A) CsgA CM und CsgA CD present CsgA constructs with additional edge strands (R1c and R5c shown in light blue), Bar-$(CsgA)_n$-Bar and Sumo-$(CsgA)_n$-Sumo present CsgA constructs fused to barnase (green) or sumo (yellow). CsgA repeating units (R1 – R5) are shown in blue, protease cleavage sites, linker peptides and affinity tags are highlighted. **(B)** Residues chosen for proline/asparagine substitution. Residues substituted by proline or lysine were highlighted in orange, red indicates residues, substituted by lysines and yellow indicates residues, substituted by proline.

CsgA CM (capped monomer) and CsgA CD (capped dimer) were further modified via site directed mutagenesis. In each case, one to two prolines or lysines were introduced in R5 cap, named R5c, and R1 cap, named R1c, of the same protein construct. As the positions of inward or outward pointing residues and hydrogen bond donors or acceptors were not known at this time point, odd and even numbered residues were chosen for modifications of each edge strand (**Figure 4-26 B**). Since the effect of charged side-chains depends on inward-pointing charges, it was important to substitute inward-pointing residues with lysine. Proline disfavours further β-strand interactions due to a missing hydrogen bond donor and a very strong local twist (Richardson and Richardson, 2002).

To enforce the effect of proline and lysine, each edge strand was modified twice in addition to single modifications of the strands. Double variants comprised variants with one modification in R5c and a modification in R1c at the corresponding position. Quadruple variants display in each case two modifications in R5c and R1c. Altogether 24 proline and lysine variants (**Table 3-7**) were generated and analysed. Barnase and sumostar were fused without further modifications. Short linker peptides were used to avoid undesired flexibility and thus a missing masking effect.

The expression of plasmid constructs, which correspond to *Ec*CsgA fusion proteins, were first analysed in order to differ soluble and insoluble expression products. For proline and lysine variants of CsgA CM, CsgA CD as well as Bar-$(CsgA)_2$-Bar and Sumo-$(CsgA)_2$-Sumo no soluble protein was obtained. In contrast to this insoluble expression, Bar-CsgA-Bar and Sumo-CsgA-Sumo were partly soluble. Short induction times and a low induction temperature (20 °C) resulted in soluble expression products as shown in **Figure 4-27**.

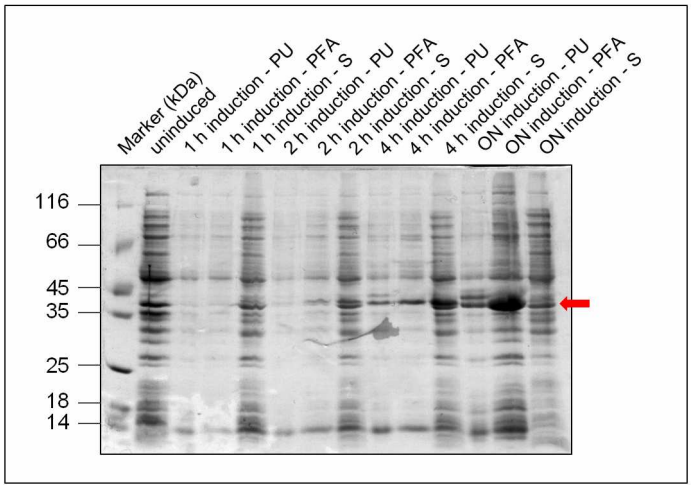

Figure 4-27: Expression test analysis of Bar-CsgA-Bar. 12 % SDS-PAGE (Coomassie staining) according to expression test analysis of Bar-CsgA-Bar. *E. coli* BL21 cells were transformed with pET 11d Bar-CsgA-Bar and induced with 250 µM IPTG. Cell growth and induction were quenched at suitable time points and cells were further processed using BugBuster reagent. Since *Ec*CsgA fibrils are not readily soluble in common denaturants, pellet fractions were solubilised with urea (PU) or formic acid (PFA) to characterise the aggregates for *Ec*CsgA amyloidogenic behaviour. Supernatants (S) were directly loaded onto the gel.

4.4.2 Protein Production and Purification

Similar to other *Ec*CsgA constructs that were analysed during this study, the fusion variants were produced as histidine-tagged proteins.

Variants of CsgA CM, CsgA CD as well as Bar-$(CsgA)_2$-Bar and Sumo-$(CsgA)_2$-Sumo were obtained in inclusion bodies and purified under denaturing conditions according to an initial inclusion body purification as illustrated in **chapter 3.4.5** and in **Figure 4-28**.

After cultivation, cell harvest and lysis, IBs were pre-purified by repetitive steps of re-suspending in appropriate washing buffers and spinning. The monomeric analogues were solubilised as described. IBs of dimeric analogues were extremely stable and thus initially solubilised in formic acid before adding the proper solubilisation buffer. Further purifications steps were identical to processing *Ec*$CsgA_{21-151}$ and *Ec*$CsgA_{40-151}$ (**chapter 4.3.3**).

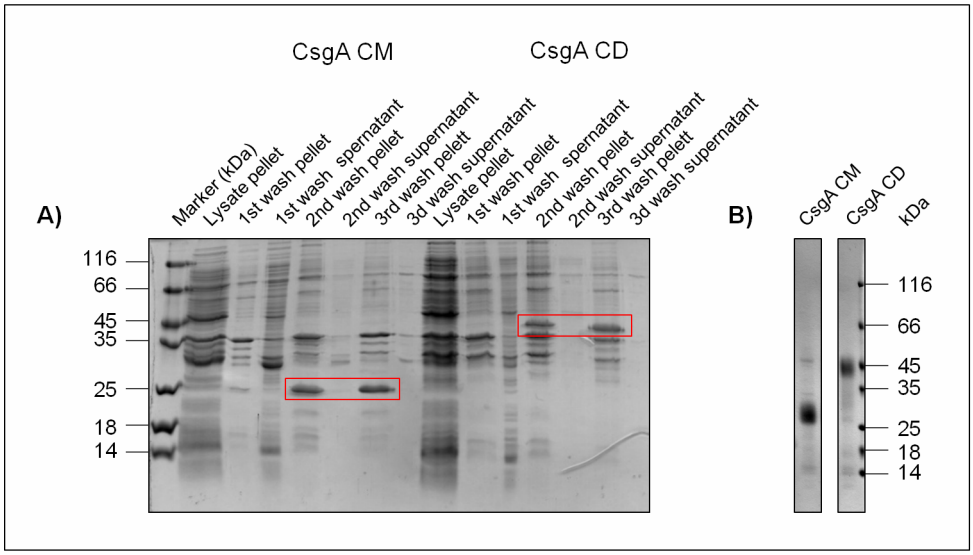

Figure 4-28: Purification of CsgA CM and CsgA CD variants. (A) 15 % SDS-PAGE (Coomassie staining) of initial purification of CsgA CM and CsgA CD inclusion bodies. 1st wash includes IB wash buffer A, 2nd and 3rd wash include lysis buffer. **(B)** 4-12 % NuPAGE® (Coomassie staining) of CsgA CM and CsgA CD according to IMAC.

In contrast to the insoluble expression products, barnase- and sumo-constructs that contain only one copy of the *Ec*CsgA amyloid core (Bar-CsgA-Bar and Sumo-CsgA-Sumo) were purified natively as described in **chapter 3.4.6**. The yield of soluble fusion proteins was increased by a cell lysis protocol with a combination of lysozyme and suitable detergents instead of sonication. After this gentle cell disruption, the soluble proteins were directly processed via IMAC under native conditions resulting already in quite pure protein fractions (**Figure 4-29**), which were applied to further experiments.

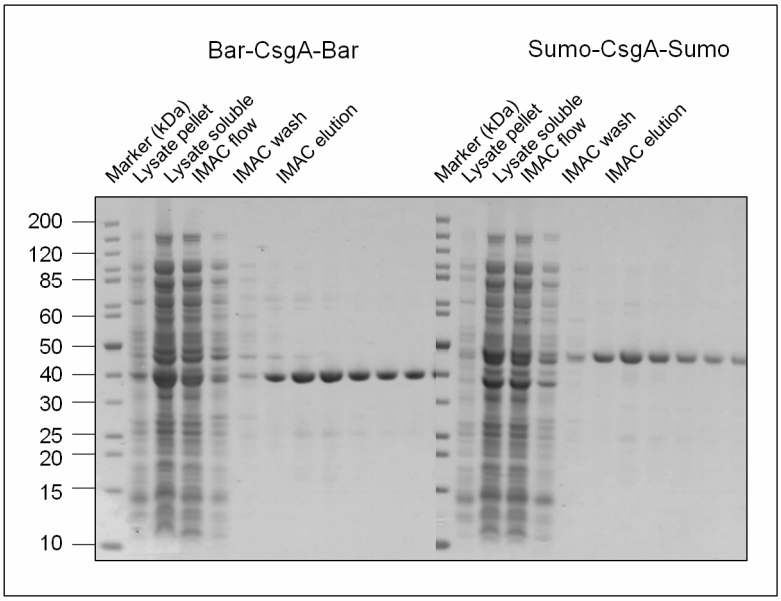

Figure 4-29: Purification of soluble Bar-CsgA-Bar and Sumo-CsgA-Sumo. 4-12 % NuPAGE® (Coomassie staining) of production and isolation of Bar-CsgA-Bar and Sumo-CsgA-Sumo.

4.4.3 Insoluble Expression Products Behave like *Ec*CsgA wt

Insoluble expression products of CsgA CM and CsgA CD variants were analysed further in order to study their amyloidogenic behaviour. Incubation of the purified proteins in native buffer after desalting resulted in protein aggregation. Suitable biophysical methods were employed to analyse the amyloid character of these aggregates as illustrated below.

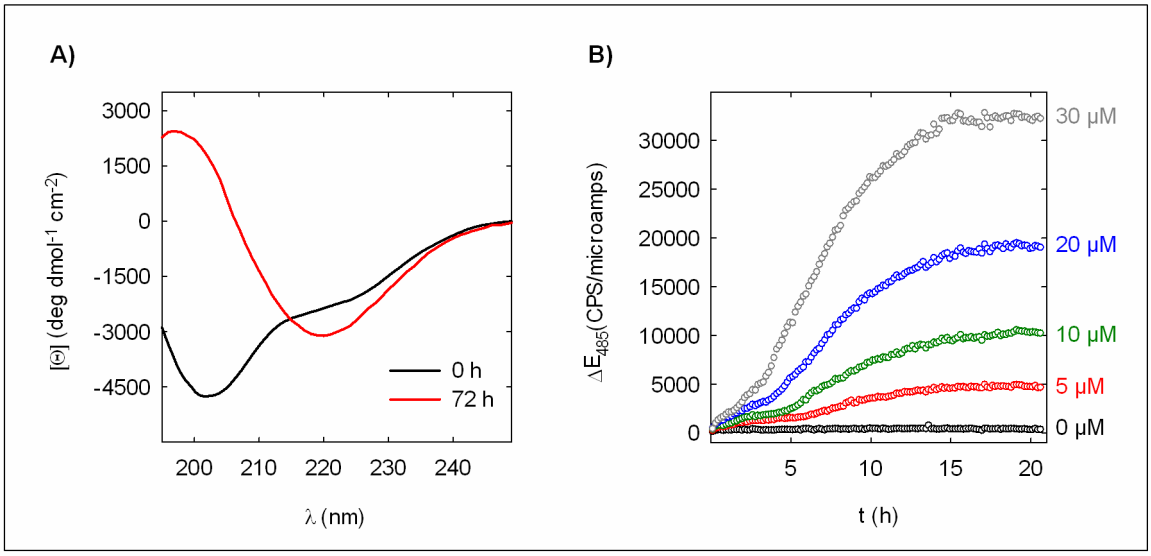

Figure 4-30: Biophysical characterisation of CsgA CM N26K/N162K. (A) Far-UV-CD spectra of CsgA CM N26K/N162K after desalting (black) with 50 mM KPi pH 7.2 and 3 days incubation (red). 20 spectra were accumulated using a cuvette with 1 mm path length. **(B)** ThT fluorescence depending on time and protein concentration. 50 µM ThT were mixed with freshly desalted CsgA CM N26K/N162K. Fluorescence was measured every 10 min at 485 nm after excitation at 442 nm.

Similar to *Ec*CsgA$_{21-151}$ (**chapter 4.3.4)**, CD spectra of the formed aggregates revealed the typical minimum around 218 nm and maximum around 195 nm as well. The CD spectra, which were recorded immediately after desalting, revealed a strong minimum around 200 nm indicating the random coil conformation. Thus, CD spectra of protein samples directly after buffer exchange and incubation under native conditions (**Figure 4-30 A**) confirmed the transition into a β-sheet rich fold (Ranjbar and Gill, 2009).

ThT mixing experiments displayed a strong ThT fluorescence and distinct phases during polymerisation illustrated in **Figure 4-30 B**. The appearance of concentration dependent lag phases was congruent to former studies of *Ec*CsgA$_{21-151}$ and *Ec*CsgA$_{40-151}$ (**chapter 4.3.4**) and other amyloidogenic proteins (Lomakin *et al.*, 1997; Wang *et al.*, 2010).

Finally, fibrillar structures were obtained by electron microscopy shown in **Figure 4-31**.

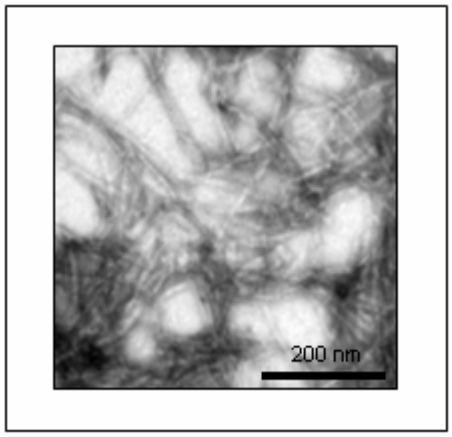

Figure 4-31: Electron micrograph of aggregates formed by CsgA CM N26K/N162K.

As shown for CsgA CM N26K/N162K, the results of the biophysical characterisation of CsgA CM and CsgA CD fusion proteins are consistent with the behaviour of typical amyloids outlined in **chapter 1.1.2**.

In contrast to the intent of generating soluble analogues of *Ec*CsgA oligomers, both proline and lysine variants of CsgA CM and CsgA CD fusion proteins seem to assemble like *Ec*CsgA wt into specific amyloid aggregates.

4.4.4 Folding Behaviour of Soluble Expression Products

The fusion proteins Bar-CsgA-Bar and Sumo-CsgA-Sumo were obtained in the soluble form and purified under native conditions. Therefore, the question about the stability and the folding level especially of the *Ec*CsgA was addressed. In order to analyse the folding

behaviour of these fusion variants, the proteins were subjected to size exclusion chromatography, solubility tests, CD experiments and 2D-NMR.

Both proteins feature a molecular mass around ~ 37 kDa (**Table 3-14**). This suggested similar elution profiles in size exclusion chromatography experiments in case of same oligomerisation level and a globular fold. Size exclusion chromatography experiments were performed as described in **chapter 3.4.6** and the results are summarised in **Figure 4-32**. The symmetric single peak indicated homogeneous protein samples with only one oligomeric species per sample. But the elution profile of both protein variants revealed different elution volumes suggesting differences in oligomerisation or fold of the proteins.

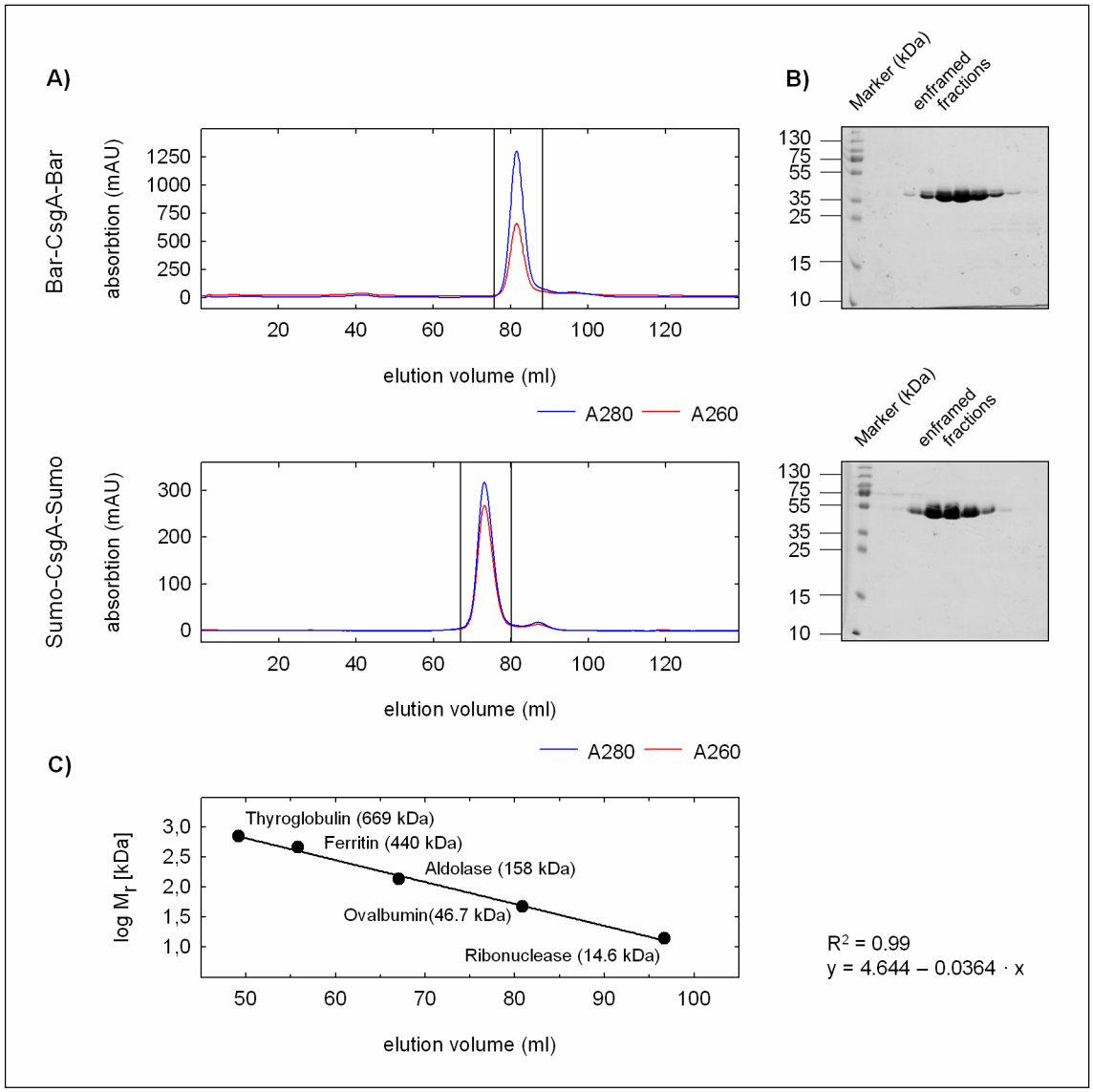

Figure 4-32: Size exclusion chromatography of Bar-CsgA-Bar and Sumo-CsgA-Sumo. (A) Elution profiles of corresponding experiments using Superdex200 (16/60). Elution occurred via 50 mM Tris/HCl pH 7.5, 150 mM NaCl with a flow rate of 1.5 mL/min. **(B)** 12 % SDS PAGE (Coomassie staining) of protein containing fractions (enframed in A) according to size exclusion chromatography. **(C)** A calibration run of Superdex200 (16/60) was performed using different proteins.

Next, the aggregation tendency of these two fusion proteins was checked by solubility tests. In low concentrated solutions (~ 0.2 mg/mL), both proteins were soluble up to several weeks at 4 °C. In concentrated samples (> 1.5 mg/mL) the variants formed aggregates at 4 °C within few hours. Contrary to the specific fibrils of *Ec*CsgA wt, these aggregates were readily soluble in 8 M urea and 0.1 % (w/v) SDS. Thus, they seem to be unspecific.

The soluble and aggregated protein fractions were analysed by CD experiments to test their secondary structural content (**Figure 4-33**). In the soluble state, CD spectra of Bar-CsgA-Bar revealed a double minimum at 228 and 208 nm and a maximum at 195 nm, typical for α-helical proteins (Ranjbar and Gill, 2009). Self-contained barnase features 21 % α-helical and 22 % β-sheet secondary structures (Baudet and Janin, 1991). CD spectra of Sumo-CsgA-Sumo displayed a double minimum at 225 and 205 nm and a maximum at 190 nm. Since the minimum at 205 nm is highly developed and CD spectra of irregular structured proteins display an intense negative Cotton effect near 200 nm (Ranjbar and Gill, 2009), this fact indicates a high random coil content. Separate sumo features 16 % α-helical and 25 % β-sheet secondary structures (Sheng and Liao, 2002). The CD spectra of the aggregates of both Bar-CsgA-Bar and Sumo-CsgA-Sumo revealed a minimum around 218 nm and a maximum around 195 nm, typical for β-sheet rich proteins (Ranjbar and Gill, 2009).

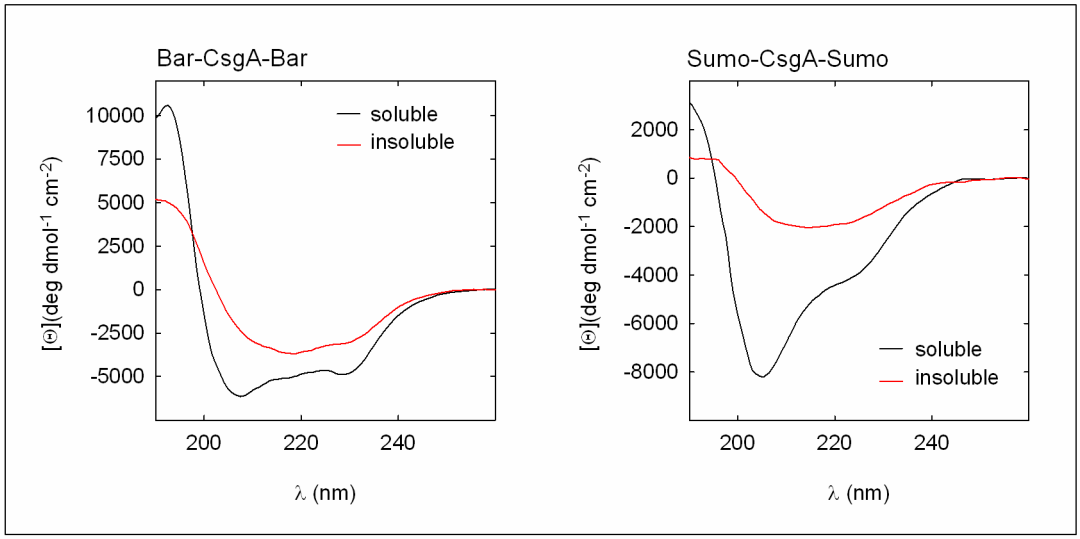

Figure 4-33: Secondary structure analysis of soluble and aggregated Bar-CsgA-Bar and Sumo-CsgA-Sumo. Far-UV-CD spectra of soluble and aggregated fusion proteins in 50 mM KPi pH 7.2. 20 spectra were accumulated using a cuvette with 1 mm path length.

Due to the promising CD spectra of soluble Bar-CsgA-Bar, this fusion protein was further investigated by 2D NMR spectroscopy to gain more information about the structural organisation. [^{15}N,^{1}H] ctHSQC correlation spectra of ^{15}N labelled Bar-CsgA-Bar were recorded accordingly. The obtained spectrum (**Figure 4-34**) revealed a remarkable number of sharp and more intense peaks in the random coil chemical shift region around 8 ppm. This fact potentially indicates numerous conformationally disordered residues.

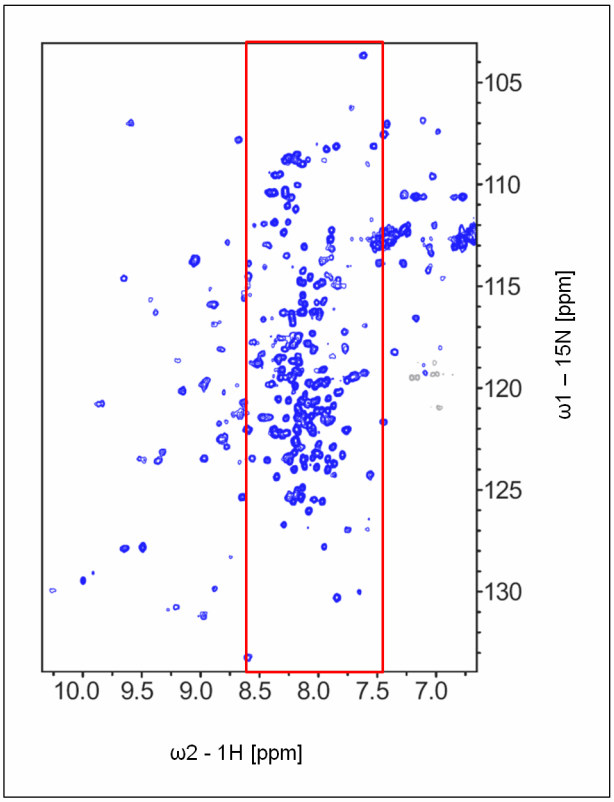

Figure 4-34: [^{15}N, ^{1}H] ctHSQC spectrum of Bar-CsgA-Bar. Spectra were recorded of 45 µM ^{15}N labelled Bar-CsgA-Bar in 50 mM KPi pH 6.

In summary, these data suggested that the fusion proteins Bar-CsgA-Bar and Sumo-CsgA-Sumo are soluble and folded to a certain extent, but not quite stable at higher concentrations. Especially the CD and NMR data of the soluble proteins indicated a high content of irregular structured and flexible residues. These data are congruent to similar experiments with the Ure2p prion domain fused to barnase (Baxa *et al.*, 2002). The prion domain was unfolded in the soluble state of the corresponding fusion protein. Thus, it is conceivable that the *Ec*CsgA moiety is not well folded in the soluble state as well. However, since Bar-CsgA-Bar and Sumo-CsgA-Sumo are soluble after expression and the formed aggregates are unspecific, the fusion proteins successfully alter the *Ec*CsgA amyloidogenic behaviour.

5 Discussion

5.1 H/D Exchange NMR Is Capable to Explain HET-s Species Barrier Breaching

In order to investigate the phenomenon of HET-s species barrier breaching, structural analysis of *Fg*HET-s$_{218-289}$ with H/D exchange NMR spectroscopy was done as part of a structure-function analysis of the HET-s homologues in collaboration with Beat Meier lab (ETH Zürich) and Sven Saupe lab (Université de Bordeaux). Thus, the following discussion takes results of all three groups into account (Wasmer *et al.*, 2010).

5.1.1 H/D Exchange Explains Remarkable Cross-Seeding

The obtained H/D exchange data of the prion forming domain of the HET-s homologues from *Fusarium graminearum Fg*HET-s$_{218-289}$ revealed four highly protected segments per molecule suggesting β-sheet elements at the positions E223-E234, A237-F245, E259-T270 and S273-G282. The corresponding residues are located in the same regions as described for *Pa*HET-s$_{218-289}$ (Ritter *et al.*, 2005) illustrated in **Figure 5-1**. Thus, both homologues seem to feature a very similar hydrogen bonding pattern.

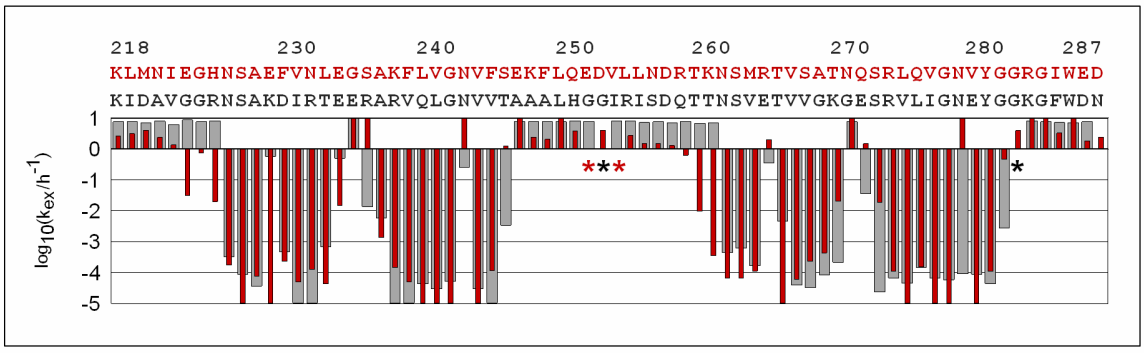

Figure 5-1: H/D exchange rates of *Fg*HET-s$_{218-289}$ and *Pa*HET-s$_{218-289}$. Red bars show the exchange rates of *Fg*HET-s$_{218-289}$ k_{ex} (h^{-1}) and grey bars indicate rates of *Pa*HET-s$_{218-289}$ (Ritter *et al.*, 2005).

In addition to *Pa*HET-s$_{218-289}$, *Fg*HET-s$_{218-289}$ features six more protected residues comprising E223-H225 and R259-K261. Since the slow exchange rates of the protected segments are monoexponential and in the same range as demonstrated for *Pa*HET-s$_{218-289}$, both homologues demonstrate an extraordinary homogeneity and stability. Fast exchange

rates within highly protected regions identified structural disturbances at positions G235-S236 and N271-Q272, as well as at the positions N243, R265 and N279 similar to *Pa*HET-$s_{218-289}$ indicating β-turns and β-arcs. In contrast to *Pa*HET-$_{s218-289}$, residues 236, 272 and 279 are less protected in *Fg*HET-$s_{218-289}$, while residue E229 was highly protected. These observations denote a different hydrogen bonding pattern at the corresponding positions.

In summary, the H/D exchange analysis of *Fg*HET-$s_{218-289}$ in comparison with *Pa*HET-$s_{218-289}$ revealed structural similarities and differences of the two homologues. Since the results are congruent, the H/D exchange analysis of *Fg*HET-$s_{218-289}$ identified all structure elements in *Fg*HET-$s_{218-289}$ that constitute the amyloid core of *Pa*HET-$s_{218-289}$. Both homologues feature β-strands and β-turns at similar positions. The protected segments 1, 3 and 2, 4 as well as the turn regions are separated in each case by 36 residues indicating a repetitive structural character. Thus, the H/D exchange NMR analysis revealed that both homologues feature very similar structures despite the low sequence identity. And this structural relationship explains the HET-s cross-seeding phenomenon, since cross-seeding presupposes structures that are highly similar to act as structural template for each other (Prusiner, 1998).

Furthermore, the comparison of the H/D exchange data enabled already detailed insights in the *Fg*HET-$s_{218-289}$ structure, since the H/D exchange data can be correlated to the high resolution solid state NMR data of *Pa*HET-$s_{218-289}$ (Van Melckebeke *et al.*, 2010; Wasmer *et al.*, 2008). *Pa*HET-$s_{218-289}$ was identified to constitute a β-solenoid structure (Wasmer *et al.*, 2008). The two pseudo-repeats of *Pa*HET-$s_{218-289}$ are suggested to form parallel β-sheets (Ritter *et al.*, 2005) and every proposed β-strand is subdivided into two shorter fragments (e.g., β1 into β1a and β1b) by β-arcs. Since the H/D exchange NMR analysis of the homologues revealed ordered and flexible residues at congruent positions, *Fg*HET-$_{s218-289}$ is capable to arrange β-strands in the same manner. Two asparagine ladders (N226/N262 and N243/N279) stabilise the amyloid structure of *Pa*HET-$s_{218-289}$. These residues are conserved among the two homologues enabling two asparagine ladders in *Fg*HET-$s_{218-289}$, too. Evidently, there are additional rigid residues in *Fg*HET-$s_{218-289}$ that comprise residues E223-H225 and R259-K261, which could form a very short β-sheet. E223 and R259 are oppositely charged and separated by 36 residues. Thus, they might form a salt bridge that stabilises the β-sheet. *Pa*HET-$s_{218-289}$ features salt bridges at positions K229/R265, E234/K270 and R236/E272. The opposite charges of the salt bridge between the residues 229 and 265 are rearranged in *Fg*HET-$s_{218-289}$ (E229/R265) and

missing charges at positions 236, 270 and 272 disable salt bridges at E234/T270 and S236/Q272. Thus, there are obvious differences in the structural and stabilising elements, too. However, both homologues feature adequate, stable structure elements at appropriate positions to form a robust and very similar fold despite these differences. The *Fg*HET-s$_{218-289}$ amyloid core is highly suggested to form a β-solenoid structure with two windings per molecule and a triangular hydrophobic core illustrated in **Figure 5-2**.

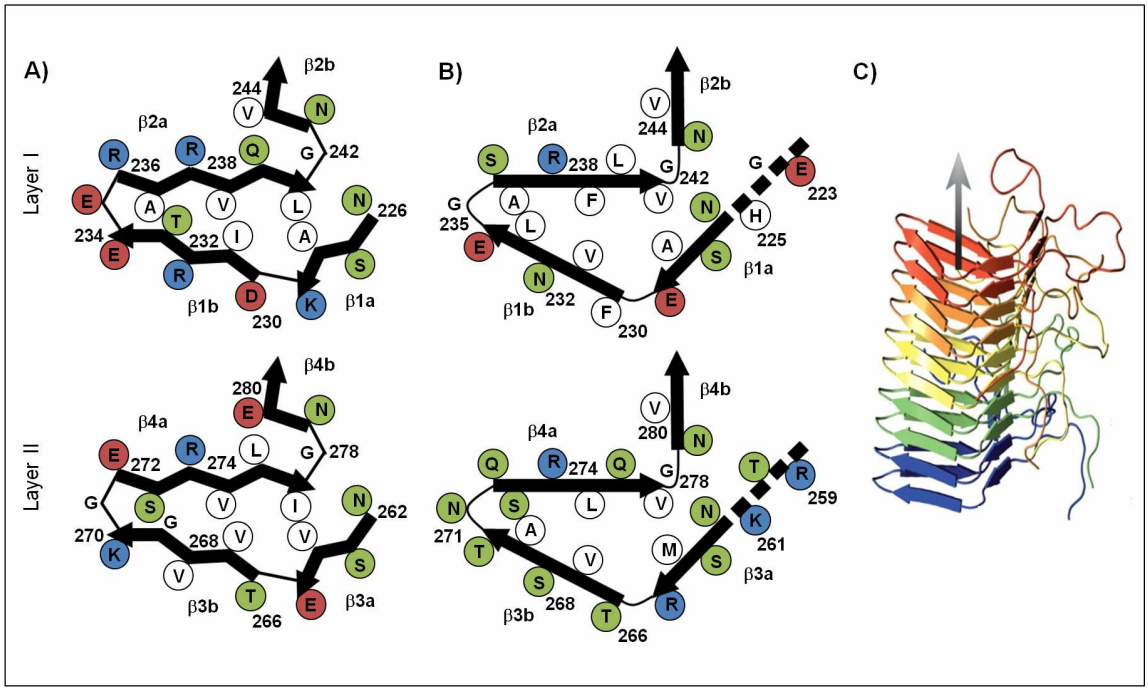

Figure 5-2: Schematic representation of the two layers of the amyloid core formed by (A) PaHET-s218-289 (structure) and (B) FgHET-s218-289 (model). Charged residues are shown in red or blue, polar but uncharged resides are highlighted in green, unpolar residues in white circles. **(C) Structure of the *Pa*HET-s$_{218-289}$ fibrils.** Side view of five stacked *Pa*HET-s$_{218-289}$ molecules (Wasmer *et al.*, 2008). Strands display in same colour correspond to one molecule, the fibril axis is indicated by an arrow.

Intermolecular hydrogen bonding via the edge strands enables the monomers to assemble along the fibril axis as shown for a *Pa*HET-s$_{218-289}$ pentamer that was calculated from NMR restraints (Wasmer *et al.*, 2008). The arrangement of hydrophobic and hydrophilic side chains is consistent to that of *Pa*HET-s$_{218-289}$. The hydrophobic core is formed by one polar and eleven non-polar residues. Outside pointing side chains are mostly polar or charged tightening the structure model for amyloid fibrils formed by *Fg*HET-s$_{218-289}$.

5.1.2 Solid State NMR Results Are Complementary and Reveal Further Structural Details

In addition to the presented H/D exchange study, solid state NMR experiments were performed by Christian Wasmer and Beat Meier (ETH Zürich) to identify residues that are involved in the β-sheet structure of *Fg*HET-s$_{218-289}$. The segments E223-S246 and D258-Y281 as well as W287 were specifically assigned. The strong dependence of the chemical shifts on the backbone amide conformation and the difference of the observed C^{α} and C^{β} secondary chemical shifts $\Delta\delta_{C\alpha}$ - $\Delta\delta_{C\beta}$ enabled information about the positions of secondary structural elements as well. α-helical and β-sheet conformations are indicated by positive and negative values respectively.

Almost all assigned residues, which displayed slow exchange rates during H/D exchange, were allocated to β-strands by solid state NMR illustrated in **Figure 5-3**.

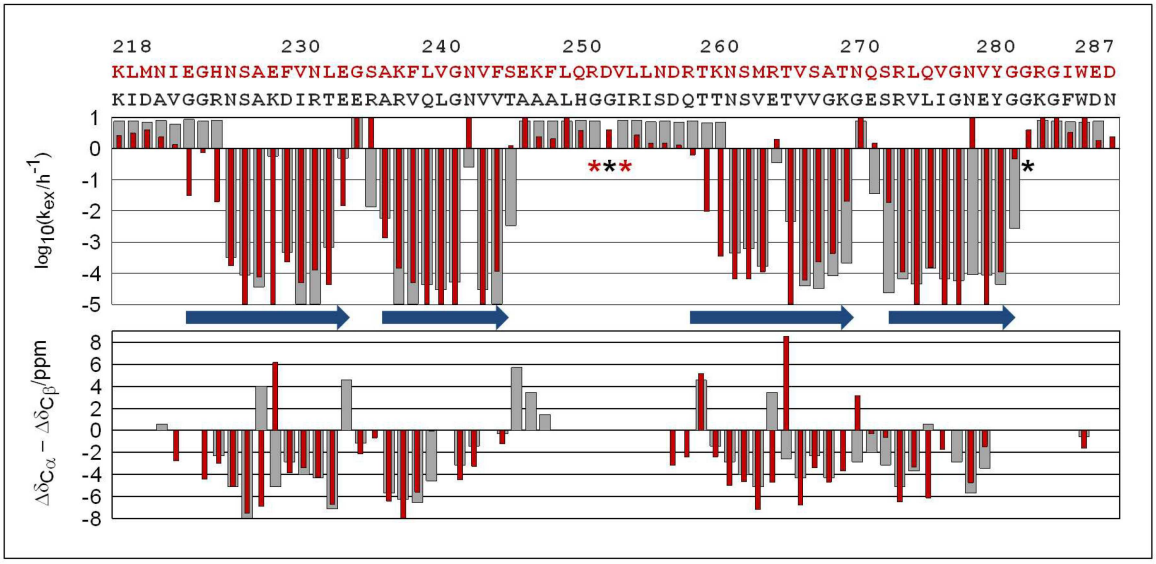

Figure 5-3: Sequence specific determination of secondary structure elements of *Fg*HET-s$_{218-289}$ and *Pa*HET-s$_{218-289}$. Upper bars indicate hydrogen exchange rates k_{ex} (h^{-1}) of *Fg*HET-s$_{218-289}$ (red) and *Pa*HET-s$_{218-289}$ (grey). Residues marked with an asterisk were detached from analysis due to missing assignment and overlap respectively. Blue arrows indicate assigned β-strands. Lower bars show the difference of Cα and Cβ secondary chemical shifts, which were extracted from solid state NMR spectra, for residues with both Cα and Cβ resonances assigned (adapted from Wasmer *et al.*, 2010).

Owing to the positive $\Delta\delta_{C\alpha}$ - $\Delta\delta_{C\beta}$ value, a structural disturbance at positions N271-S272 was identified too by the secondary shift deviation. In contrast to former studies of *Pa*HET-s$_{218-289}$, negative $\Delta\delta_{C\alpha}$ - $\Delta\delta_{C\beta}$ values were detected for presumable β-arc positions E229 and R265, while the H/D exchange data indicated a structural disturbance at position 265. In spite of slow H/D exchange rates, positive values were detected for the adjacent

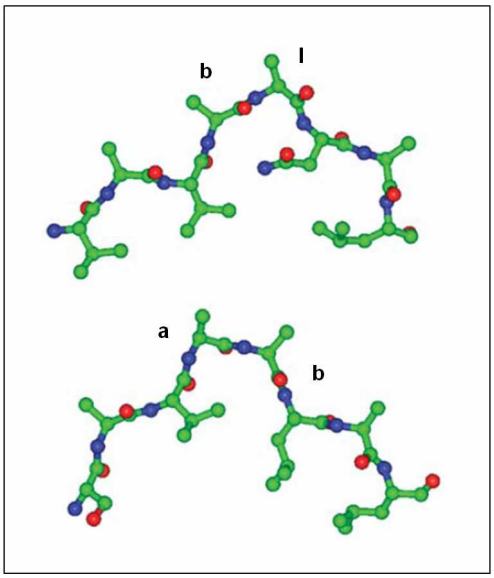

Figure 5-4: Typical two-residue β-arcs. Stick and ball representation of *bl* and *ab* arcs. Oxygen atoms are shown in red, nitrogen in blue and carbon in green (Hennetin *et al.*, 2006).

residues F230 and T266. The position of the expected secondary shift deviation is shifted by one residue.

Since there are multiple types of two-residue β-arcs (Hennetin *et al.*, 2006), *Fg*HET-s$_{218-289}$ could feature *bl* arcs, while *ab* arcs are present at the corresponding positions in *Pa*HET-s$_{218-289}$. These arrangements would show the same side-chain topology but different backbone angles (**Figure 5-4**) that probably induce differences in the secondary chemical shifts (Wishart and Nip, 1998).

In summary, solid state NMR and H/D exchange experiments complemented each other. The H/D exchange NMR analysis was complete and revealed significant information about the hydrogen bonding pattern of all *Fg*HET-s$_{218-289}$ backbone amides, while the solid state NMR analysis revealed insights in the backbone conformation of the assigned residues. Thus, both techniques identified structural similarities as well as differences of the homologous proteins and complementary specified all structure elements to form a highly similar amyloid core.

5.1.3 The β-Solenoid Fold of HET-s Is Evolutionary Conserved

The sequence specific secondary structural analysis has shown that the specific aggregates formed by *Fg*HET-s$_{218-289}$ and *Pa*HET-s$_{218-289}$ are structurally highly similar, although they show low sequence identity. It is well known that globular, soluble proteins with a homology level around 30 % typically adopt similar structures (Chothia and Lesk, 1986; Ginalski, 2006). In contrast to this behaviour, amyloids have been shown to form different polymorphic forms and much different structures already due to single point mutations (Tycko *et al.*, 2009). Thus, the obtained structural details are quite remarkable.

A closer look to the sequences (**Figure 5-5**) of the homologous domains found in *Podospora anserina* and different *Fusarium* species shows that most of the conserved residues appear in the well structured amyloid core region (pseudo-repeats 222-247 and 258-283). The asparagine residues N226, N243, N262, N279, which form two

asparagine ladders along the fibril axis in *Pa*HET-s$_{218-289}$, are conserved in all homologous proteins. Since glutamine and asparagine ladders are crucial for prion aggregation (DePace *et al.*, 1998; Osherovich *et al.*, 2004; Ross *et al.*, 2005), these asparagines seem to be essential to provide the HET-s amyloid structure. The glycines G240 and G278 that enable the formation of sharp β-arcs and the glycine-rich loop at the C-terminus and W287, which are assumed to make contacts with β-strand 2 and 4 (Van Melckebeke *et al.*, 2010), are conserved, too. Several residues, which are not conserved, are substituted to homologous residues. Furthermore, the amyloid core region comprises numerous hydrophobic residues that are capable to constitute a hydrophobic core.

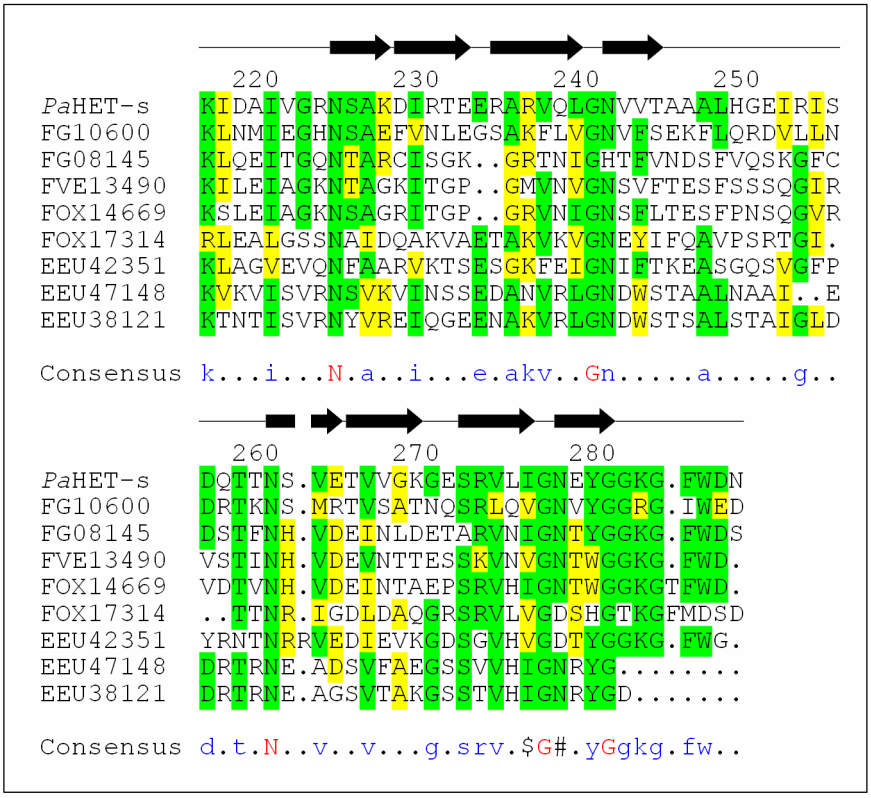

Figure 5-5: Sequence alignment of HET-s prion forming domain from *Podospora anserina* and homologues found in different *Fusarium* species. Primary structure of *Pa*HET-s$_{218-289}$ is compared to homologous domains. Sequences of the homologues correspond to GenBank accession number, FG10600 and FG08145 are from *F .graminearum*, FVE13490 from *F. verticllioides*, FOX17314 and FOX14669 are from *F. oxysporum*, EEU42351, EEU47148 and EEU38121 are from *F. solani*. Conserved residues are highlighted in green, homologous residues in yellow. In the consensus sequence conserved residues are shown in red and conserved amino acid exchanges highlighted in blue, $ indicates anyone of I or V, # anyone of D, E, N or Q. Secondary structures elements of the β-solenoid structure of *Pa*HET-s$_{218-289}$ (Wasmer *et al.*, 2008) are shown on the top of the alignment. Analysis was done with the program MultAlin (Corpet, 1988).

Thus, all residues that are crucial to form the β-solenoid structure of HET-s are highly conserved among the homologues. Similar to homologous globular proteins (Chothia and Lesk, 1986; Ginalski, 2006), these conserved residues enable highly similar structures at

reasonable levels of sequence identity. In spite of the evolutionary divergence of the HET-s sequences, this fact ensures the structural similarity of the two prion domains. The HET-s β-solenoid fold seems to be evolutionary conserved.

5.1.4 Conclusion

The presented results of the *Fg*HET-s$_{218-289}$ structure analysis clearly demonstrated the structural relationship of the two HET-s homologues (**Figure 5-6**) and thus explained the cross-seeding effect in spite of low sequence identity.

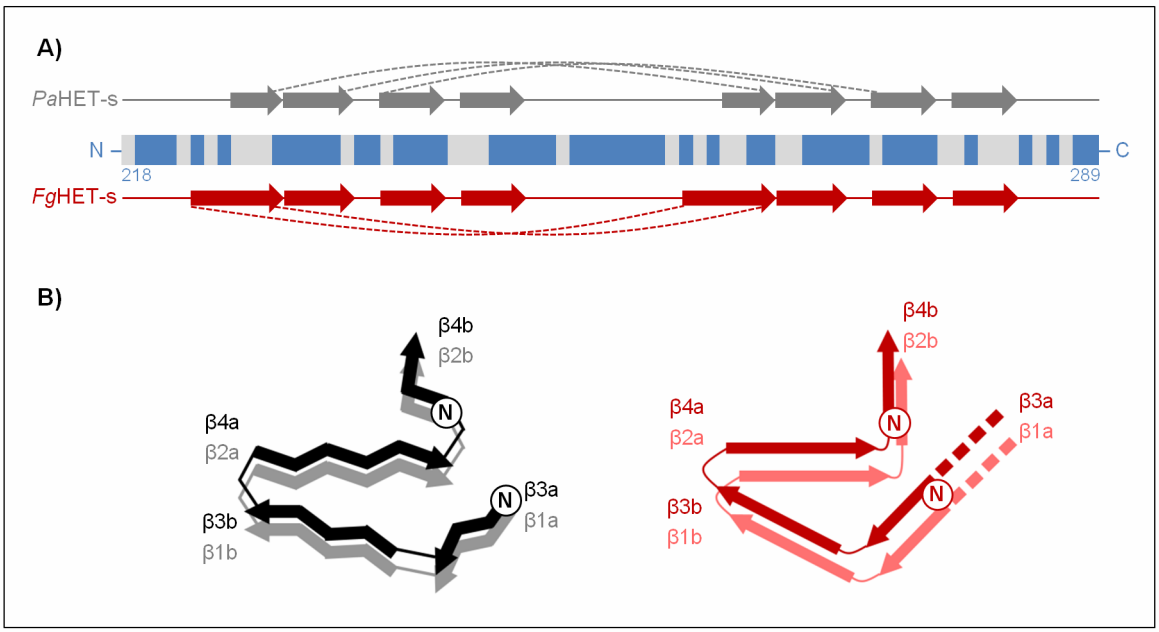

Figure 5-6: Structural relationship of the HET-s homologues. (A) Secondary structure elements of *Pa*HET-s$_{218-289}$ and *Fg*HET-s$_{218-289}$ aligned to the amino acid sequence. Arrows indicate β-strands. Furthermore, salt bridges are shown as dashed lines. The amino acid sequence of both homologues is shown as blue bar; sequence identity is shown in white. **(B)** Schematic representation of the two layers of the amyloid core formed by *Pa*HET-s218-289 (structure, black) and *Fg*HET-s218-289 (model, red) showing all β-strands, β-arcs and asparagine ladders.

The presented analysis revealed structural similarities and differences of the two HET-s homologues. Both prion domains share homogeneous and stable fibrils. They feature all major structural elements to form a highly similar β-solenoid fold. At last this structural similarity facilitates the amyloid templating in spite of low sequence identity. Similar to globular proteins, the conserved residues facilitate the highly similar structure and indicate an evolutionary conservation of the HET-s β-solenoid fold.

Since the H/D exchange NMR analysis was complete and complementary data of *Pa*HET-s$_{218-289}$ were available, the recent H/D exchange analysis facilitated the identification of structural similarities and differences of the two HET-s homologues.

Thus, all major structure elements of the HET-s fold were detected in *Fg*HET-s$_{218-289}$ fibrils without a complete and time-consuming solid state NMR analysis. Furthermore, the H/D exchange analysis enabled information about homogeneity and stability of the amyloids. Doubtless, H/D exchange NMR is a sensitive tool to compare homologous proteins with regard to their structures, homogeneity and stability and is capable to explain the HET-s cross-seeding phenomenon.

5.2 The Sophisticated Topology of the Major Curli Subunit

The structure-function relationship of CsgA is of high interest, since CsgA is the major subunit of the curli fibrils, which feature manifold functions in bacterial pathogens.
Therefore, a sequence specific analysis of the topology of *Ec*CsgA fibrils was performed within this study and the corresponding results of the H/D exchange NMR experiments and solvent-accessibility studies revealed first residue-specific information about the major curli subunit.

5.2.1 H/D Exchange NMR Experiments and Thiol-Labelling Reveal a Possible Fold of *Ec*CsgA

The H/D exchange analyses of the two variants *Ec*CsgA$_{21-151}$ and *Ec*CsgA$_{40-151}$ were congruent and complementary. Both H/D exchange experiments revealed the same structural pattern and indicated similar homogeneity and stability of the variants. Despite the lower time resolution of ctHSQC experiments, H/D exchange rates were successfully determined for 87 % of all backbone amides of mature *Ec*CsgA including almost all residues (94 %) within the amyloid core (residues 43-151). In consequence, the obtained exchange rates clearly specified the flexible N-terminus and the stable amyloid core of mature *Ec*CsgA (**Figure 5-7**). The slow exchange rates ($\leq 10^{-2}$ h^{-1}) of protected segments are congruent to the precedent analysis of *Fg*HET-s $_{218-289}$ (discussed in **chapter 5.1**) and other amyloids (Luhrs *et al.*, 2005; Ritter *et al.*, 2005; Vilar *et al.*, 2008) reflecting the extraordinary stability of *Ec*CsgA fibrils. Strong overlap of the intramolecularly conserved asparagines N54, N77, N99, N122 and N144 hampered the interpretation of their H/D exchange behaviour. The β-strand incorporation of these residues remained elusive. Furthermore, a few residues revealed unexpected biphasic exchange behaviour within protected segments discussed in **chapter 5.2.3**. However, the exchange analysis identified distinct regions of flexible and ordered segments in a repetitive manner similar to the established repertoire of secondary structural elements within amyloid fibrils outlined in **chapter 1.1.2**. The presented H/D exchange NMR analysis identified ten distinct β-strands within five imperfect sequence repeats (**Figure 5-7**). Every repetitive segment forms two β-strands, which are connected by a short β-turn. The N-terminal β-strand (β1, β3...) of each repeat comprises seven residues; the C-terminal strand (β2, β4...) constitutes of seven to eight residues depending on the incorporation of the intramolecularly conserved

asparagine. The corresponding strands are connected by an intramolecularly conserved glycine rich turn and the β-strands of adjacent repeats are linked by five to six flexible residues.

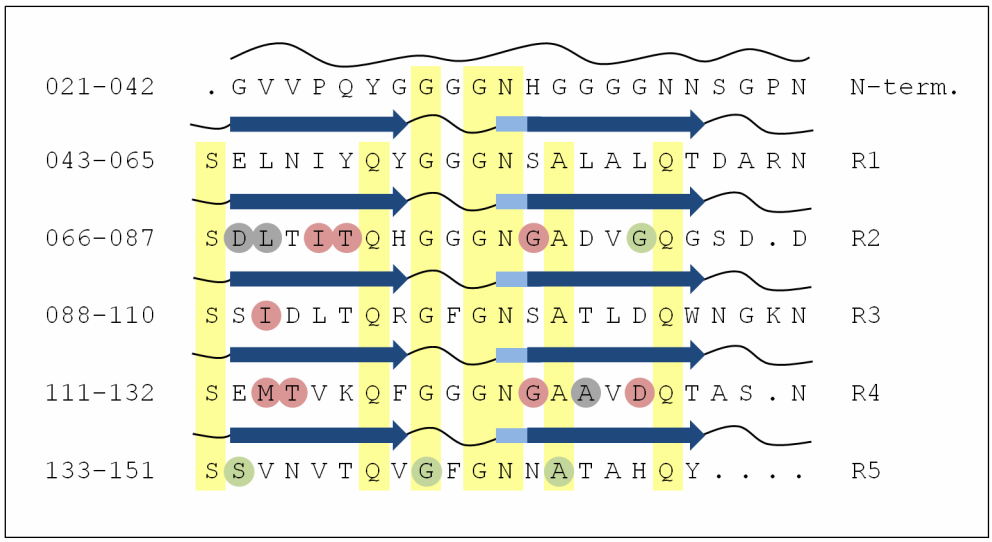

Figure 5-7: Sequence specific position of secondary structure elements within *Ec*CsgA fibrils probed by H/D exchange. Sinuous lines represent flexible regions, while blue arrows present regular segments of β-strands. The incorporation of the intramolecularly conserved asparagines remained unclear shown in light blue. Residues with biphasic H/D exchange behaviour are encircled: grey indicates residues that exchanged in a biphasic manner in both variants *Ec*CsgA$_{21\text{-}151}$ and *Ec*CsgA$_{40\text{-}151}$; green indicates biphasic exchange in E*c*CsgA$_{21\text{-}151}$, red in *Ec*CsgA$_{40\text{-}151}$. Intramolecularly conserved residues are highlighted in yellow.

Furthermore, solvent-accessibility studies (**chapter 4.3.7**) indicated that every β-strand and therefore every β-sheet of the fibril core features a solvent accessible side.

Owing to these facts, *Ec*CsgA is suggested to form a β-solenoid structure with a solvent protected core illustrated in **Figure 5-8**. This core is constituted of two β-sheets that are formed by β1, β3, β5, β7 and β9, and β2, β4, β6, β8 and β10 respectively. Since the intramolecularly conserved glutamine and asparagine residues would stack on top of each other in this structure, two glutamine-ladders and one asparagine-ladder stabilise the amyloid core as suggested before (Collinson *et al.*, 1999). Furthermore, the charged residues D127 and H149 are located in probably adjacent positions and might form a stabilising salt bridge. Similar to the HET-s solenoid (discussed in **chapter 5.1.3**), the distribution of polar and hydrophobic residues along the sequence enables a hydrophobic core. The proposed core of every coil is formed by four hydrophobic residues and the stacking asparagine and glutamine residues. Outside pointing side chains are mostly polar or charged.

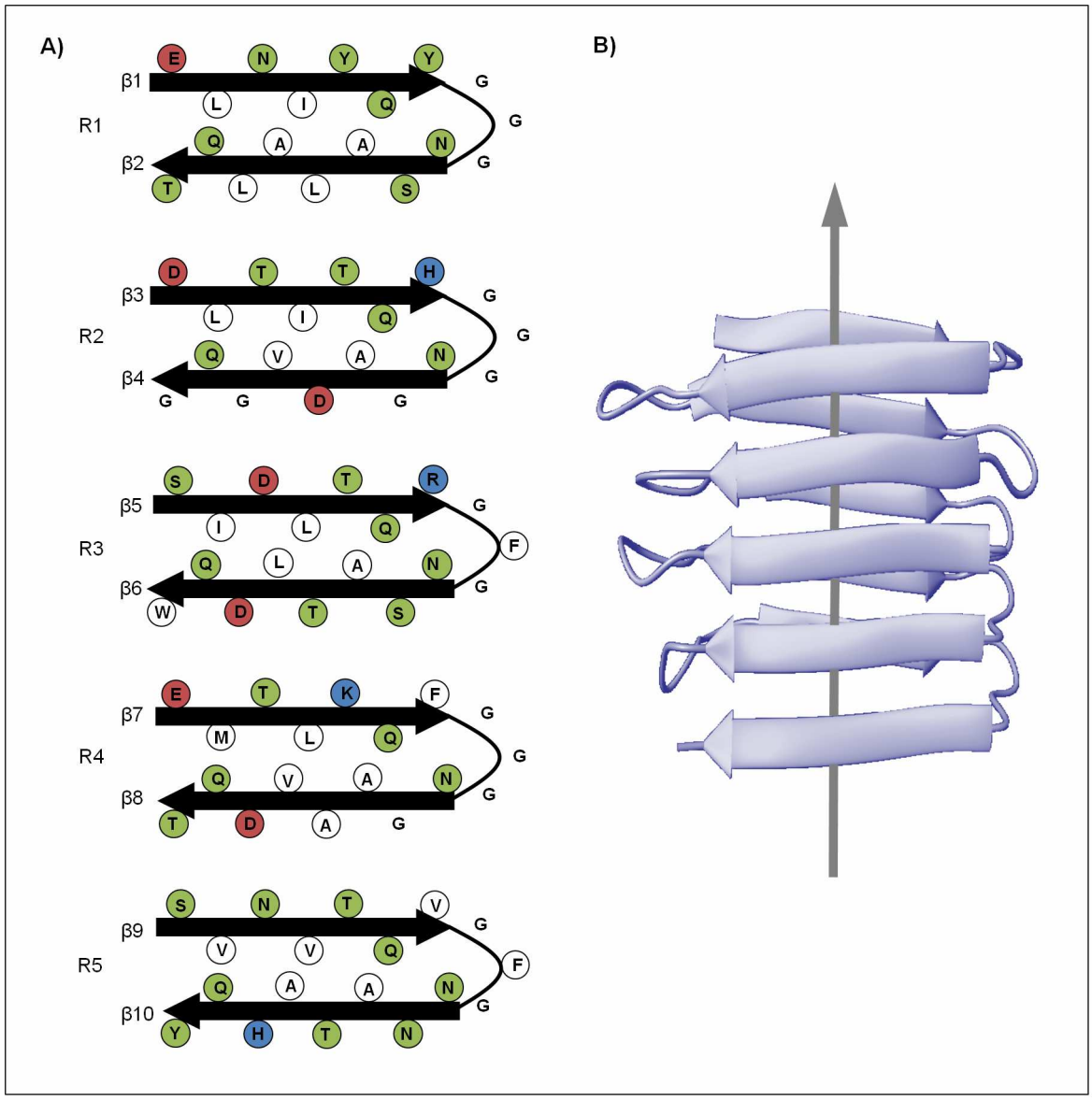

Figure 5-8: Schematic representation of the five repetitive layers and the proposed fold of the amyloid core formed by mature *Ec*CsgA. (A) The repetitive layers each form a strand-turn-strand motif. Charged residues are shown in red or blue, polar but uncharged residues in green and unpolar residues in white circles. **(B)** The model of the proposed fold is shown in a ribbon diagram.

In this proposed fold, each intramolecularly conserved alanine constitutes the hydrophobic core with three other hydrophobic residues congruent to the sequence alignment. However, the H/D exchange NMR analysis did not exclude a kink or a β-arc at these positions due to the probable fast H/D exchange behaviour of these residues. On the assumption that the intramolecularly conserved asparagines are involved in backbone hydrogen bonding, a short β-strand prior to the alanines is conceivable. The corresponding β-sheet would be stabilised by the asparagine ladder. In order to complement the presented data and to specify the microstructure of the *Ec*CsgA amyloid core, a complete solid-state NMR analysis is necessary Furthermore, stable cysteine variants, which were identified to carry a

solvent accessible cysteine, should be used to prove intra- and intermolecular contacts by fluorescence excimer studies.

In summary, the recent study indicates that the major curli subunit CsgA forms a β-solenoid structure with five windings per molecule illustrated in **Figure 5-8**. The proposed fold agrees with the observed H/D exchange rates, side-chain solvent accessibility, the distribution of polar and hydrophilic residues along the sequence and the repetitive sequence character. Intramolecular hydrogen bonding via the edge strands enables the monomers to polymerise along the fibril axis. A head to tail arrangement would facilitate the formation of stabilising glutamine and asparagine ladders along this axis. Furthermore, the formation of a β-solenoid structure is consistent to former solid state NMR data, which excluded an in-register parallel β-sheet architecture of *Ec*CsgA (Shewmaker *et al.*, 2009).

5.2.2 The Repetitive Primary Structure of *Ec*CsgA Enables a β-Solenoid Fold

The presented data suggested *Ec*CsgA to form a β-solenoid similar as the HET-s homologues (discussed above; Wasmer *et al.*, 2008; Wasmer *et al.*, 2010). Solenoids are elongated structures of repeating β-strand units that coil around the solenoid axis (Kajava and Steven, 2006). A closer look at the sequence of *Ec*CsgA and the positions of regular secondary structural elements (**Figure 5-9**) already reveals a characteristic amino acid composition with regard to β-solenoids.

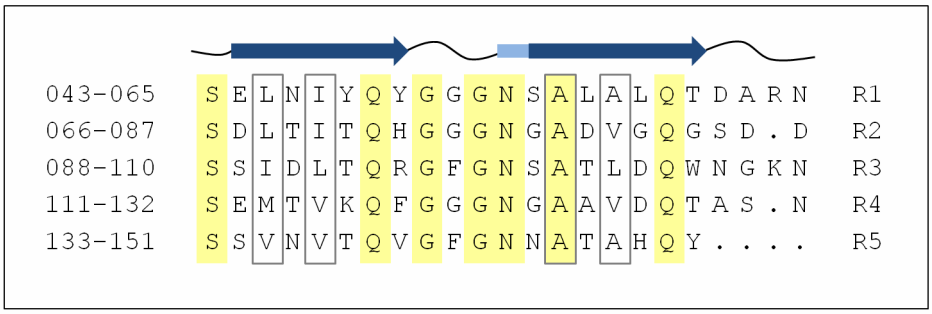

Figure 5-9: Primary and secondary structure of the *Ec*CsgA amyloid core. Comparison of the primary structure of the five repeating units and positions of secondary structural elements along the sequence. Intramolecularly conserved residues are highlighted in yellow, probably hydrophobic ladders are enframed in grey. Sinuous lines represent flexible regions, while blue arrows present regular segments of β-strands. The incorporation of the intramolecularly conserved asparagines remained unclear shown in light blue.

Similar to other β-solenoids (Kajava and Steven, 2006), *Ec*CsgA contains evident sequence repeats (Collinson *et al.*, 1999; Hammar *et al.*, 1996). Every repeated unit shows the

consensus sequence of $Sx_5QxGxGNxAx_3Qx_{1-5}$ (Collinson *et al.*, 1999) and forms a strand-turn-strand motif. The repeats contain several xyx patterns (x denotes any hydrophobic residue, y denotes mostly polar residues), which correspond to the β-strands and enable the formation of a hydrophobic core. Turns and loops are made of glycines and polar residues similar to what has been found in other β-solenoids (Hennetin *et al.*, 2006). The *Ec*CsgA sequence includes intramolecularly conserved asparagine and glutamine residues, which are able to form stabilising ladders probably essential for amyloid aggregation (DePace *et al.*, 1998; Osherovich *et al.*, 2004; Ross *et al.*, 2005). In general, β-solenoids are often enriched with asparagines forming stabilising ladders (Yoder and Jurnak, 1995). Therefore, the *Ec*CsgA sequence obviously features fundamental elements to form a repetitive β-solenoid structure. Similar to *Ec*CsgA, HET-s consists of two pseudo-repeats enabling the β-solenoid fold as discussed in **chapter 5.1**.

In contrast to *Ec*CsgA and HET-s, several fibrils of disease associated amyloids and numerous yeast prions have been shown to feature a parallel, in-register β-sheet arrangement (Balbach *et al.*, 2002; Baxa *et al.*, 2007; Heise *et al.*, 2005; Shewmaker *et al.*, 2006; Wickner *et al.*, 2008a). Obviously, the primary structures of these proteins differ from the sequences of CsgA, HET-s and other β-solenoid proteins as illustrated for $A\beta_{1-42}$, α-synuclein and $Ure2p_{1-89}$ in **Figure 5-10**.

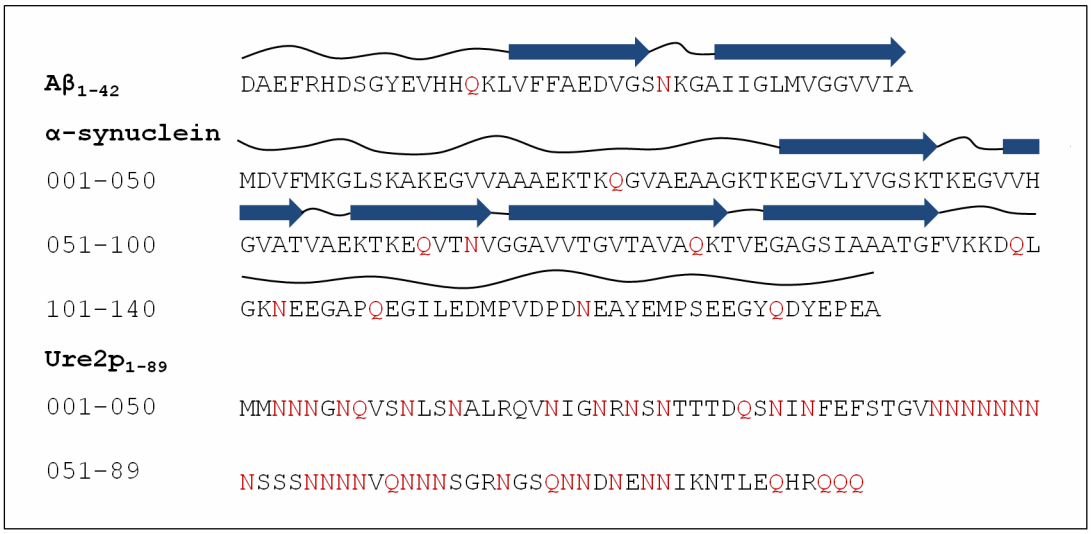

Figure 5-10: Primary and secondary structure of $A\beta_{1-42}$, α-synuclein and $Ure2p_{1-89}$. Asparagine and glutamine residues are shown in red. For $A\beta_{1-42}$ and α-synuclein the positions of secondary structural elements are shown along the sequences (Luhrs *et al.*, 2005; Vilar *et al.*, 2008). Sinuous lines present flexible regions, while blue arrows present regular segments of β-strands.

The sequences show limited repetitive character, probably stacking asparagine and glutamine residues appear in a random manner. Furthermore, the short peptide fragments

of Aβ disable repeated coils within single Aβ-molecules. In consequence, the cross-β structure of these amyloids is formed by parallel, in-register β-sheets via stacks of identical strand-turn-strand motifs (Aβ) or superpleated β-structures (Ure2p, Rnq1p, Sup35, α-synuclein) illustrated in comparison to stacks of β-solenoids in **Figure 5-11** (reviewed by Kajava *et al.*, 2010). While disease related amyloids are soluble in their native form (**chapter 1.1.3**, **Table 1-1**) and yeast prions form amyloid fibrils only at very low frequencies (Tuite and Cox, 2003), CsgA and HET-s are mostly fibrillar in the native form to fulfil their distinct functions (Chapman *et al.*, 2002; Coustou *et al.*, 1997). Similar to the HET-s β-solenoid fold (Ritter *et al.*, 2005; Wasmer *et al.*, 2010), the CsgA structure and sequence seem to be evolutionary optimised for amyloid fibril formation.

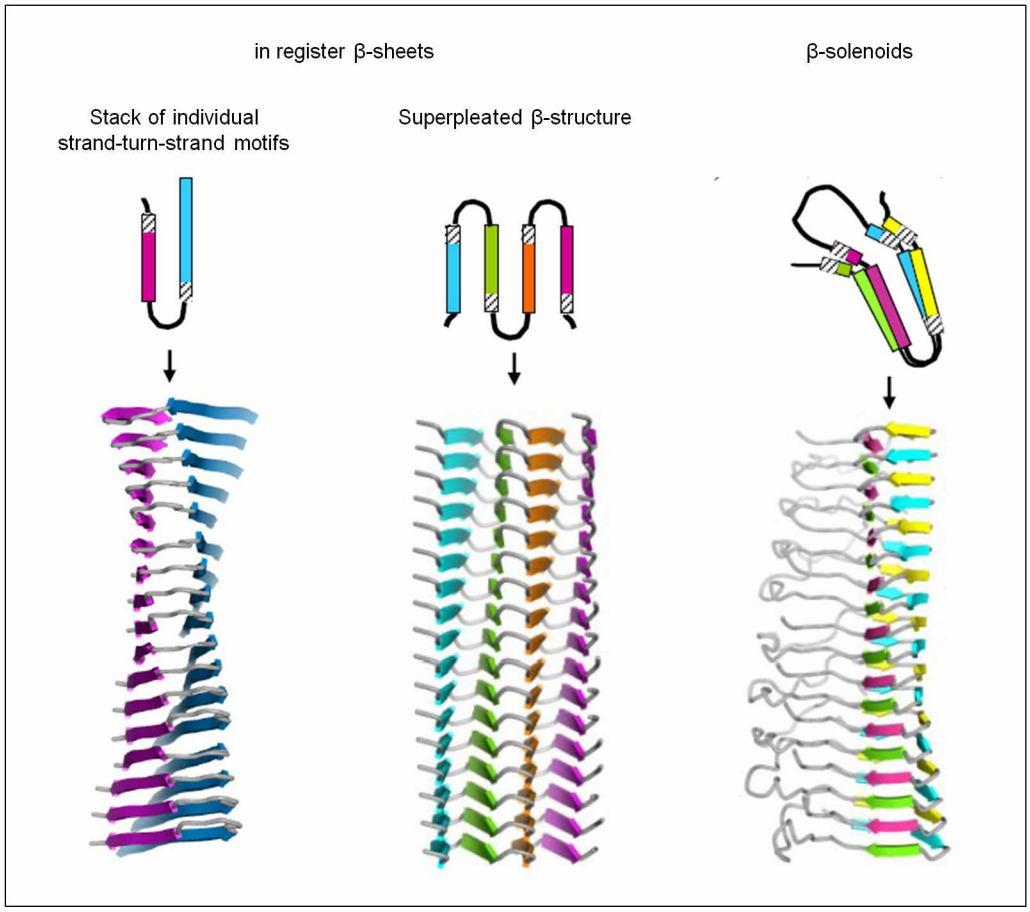

Figure 5-11: Types of cross-β structure within amyloid fibrils (original figure from Kajava *et al.*, 2010). Top: axial views of the repetitive structural units. Bottom: lateral view of cross-β structures formed by stacking of the corresponding repetitive units.

5.2.3 Biphasic *Ec*CsgA Exchange Behaviour Indicated Local Heterogeneity in *Ec*CsgA Fibrils

Illustrated above in **Figure 5-7**, the H/D exchange analysis of *Ec*$CsgA_{21-151}$ and *Ec*$CsgA_{40-151}$ revealed several backbone amides that exchanged in a biphasic manner indicating less homogeneity at these positions. A few residues featured a fast and a slow exchanging population in both, others only in one of the two analysed *Ec*CsgA variants *Ec*$CsgA_{21-151}$ and *Ec*$CsgA_{40-151}$.

Biphasic H/D exchange was also detected in former analysis of $A\beta_{1-42}$ and α-synuclein fibrils (Luhrs *et al.*, 2005; Vilar *et al.*, 2008). In contrast to the random appearance of biphasic exchange in *Ec*CsgA fibrils, distinct segments of $A\beta_{1-42}$ and α-synuclein revealed less homogeneity. In $A\beta_{1-42}$ fibrils all residues of the N-terminal flexible tail exhibited two distinct populations in H/D exchange, while the amyloid core was highly protected (Luhrs *et al.*, 2005). Contrariwise, α-synuclein fibrils featured a biphasic behaviour for all residues of the fibril core. The two subpopulations exchanged in a slow and a very slow manner. The rates of the two corresponding subpopulations vary depending on the fibril age. Thus, α-synuclein features two distinct structural entities (Vilar *et al.*, 2008). Contrary to the biphasic exchange behaviour of these disease-related amyloids, the two HET-s homologues *Pa*HET-$s_{218-289}$ and *Fg*HET-$s_{218-289}$ exclusively exchanged in monophasic manner.

In *Ec*CsgA the corresponding biphasic residues appeared in a random manner along the protected residues and the subpopulations exchanged in a fast and a very slow manner. Most of these residues were located in the central repeats R2, R3 and R4. In comparison to R1 and R5, single peptides of these segments are less amyloidogenic (Wang *et al.*, 2008b; Wang *et al.*, 2007). R2, R3 and R4 bear so called “gatekeeper” residues (Wang *et al.*, 2010). Thus, the residues G78, D80, G82, D91, D104, G123 and D127 reduce the aggregation propensity in the distinct CsgA segments and avoid uncontrolled self-assembly of CsgA. In the presented H/D exchange analysis, G78, G82, G123 and D127 were identified at least in one *Ec*CsgA variant to feature one less protected population. Therefore, it is conceivable, that the less aggregation prone repeats R2, R3 and R4 including the gatekeeper residues result in less amyloidogenic behaviour of the distinct segments within the *Ec*CsgA fibrils.

Since *in vivo* the proper fibrillisation of CsgA is nucleator dependent (Hammar *et al.*, 1996), the absence of the nucleator protein CsgB could also result *in vitro* in altered CsgA

fibrils. Therefore, heteronucleated *Ec*CsgA fibrils were analysed. In comparison to preformed *Ec*CsgA fibrils, initial experiments revealed that *in vitro* formed *Ec*CsgB fibrils are no adequate seed. A further discussion follows in **chapter 5.2.4**. In contrast to the *in vivo* fibrillisation, *in vitro Ec*CsgA was fibrillised from high concentrated protein solutions (≥ 50 µM). Thus, the heterogeneity could be also a concentration dependent effect.

The unexpected behaviour of a few *Ec*CsgA cysteine variants within side-chain solvent accessibility studies (**chapter 4.3.7**) already indicated the sensitivity of the *Ec*CsgA microstructure to modified fibrillisation properties. Since H/D exchange NMR spectroscopy is capable to specify the homogeneity and stability of the fibrils as illustrated in **chapter 4.3.6** and discussed above, it is a sensitive tool to determine the influence of such modifications to the structure and stability. In order to ensure meaningful experiments using different protein variants and different fibrillisation conditions, structure, homogeneity and stability of the corresponding proteins should be probed by H/D exchange NMR analysis.

5.2.4 Heterogeneity of Heteronucleated *Ec*CsgA Fibrils

CsgB initiates the assembly of CsgA into regular amyloid fibrils *in vivo* (Hammar *et al.*, 1996). Both proteins share 30 % sequence identity and 40 % homology (Hammar *et al.*, 1995). Since CsgB presents an amyloid like template to CsgA, it was believed that self-seeding and heteronucleation are mechanistically similar processes that lead to highly similar CsgA structures (Hammer *et al.*, 2007; Wang *et al.*, 2008b).

In contrast to the H/D exchange NMR analysis of self-seeded *Ec*CsgA fibrils (**Figure 5-7**), heteronucleated *Ec*CsgA$_{40\text{-}151}$ fibrils revealed a less regular structure and strong heterogeneity (**chapter 4.3.6**, **Figure 4-21**) indicating an altered fibril structure. In recent heteronucleation experiments, *in vitro* formed *Ec*CsgB fibrils were used to seed *Ec*CsgA fibrillisation. 10 % (w/w) were added to monomeric *Ec*CsgA$_{40\text{-}151}$ to initiate the *Ec*CsgB mediated assembly. While R1 and R5 featured the same protection pattern as in self-seeded fibrils and thus a regular structure, the central repeats exhibited less regularity and an increased heterogeneity.

Explanations for this folding behaviour are the properties of *in vitro* formed *Ec*CsgB fibrils and the different amyloidogenic behaviour of the distinct repeats within *Ec*CsgA: In recent experiments *Ec*CsgB assembles without lag phase into SDS-stable but heterogeneous

fibrils. Parallel H/D exchange NMR analysis revealed that *in vitro* only *Ec*CsgB fibrils were produced showing biphasic H/D exchange behaviour for nearly all residues of the amyloid core (Madhu Nagaraj, unpublished). Thus, the used seeds seem to feature different structural entities and different contact points for monomeric *Ec*CsgA leading into probably heterogeneous and less specific *Ec*CsgA aggregates.

The repeats R1 and R5 of *Ec*CsgA are crucial for the *Ec*CsgA assembly and they are the seeding responsive domains within this protein (Wang *et al.*, 2008b). Thus, it was not plausible that R1 and R5 featured a regular secondary structure within heteronucleated *Ec*CsgA fibrils, while the central repeats were heterogeneous und irregular structured. R1 and R5 are indispensable for *Ec*CsgA to interact with *Ec*CsgB (Wang *et al.*, 2008b). But the high seed-concentration of *Ec*CsgB fibrils, as well as the hydrophobicity of the heterogeneous *Ec*CsgB core and the monomeric *Ec*CsgA molecules enable hydrophobic interactions between *Ec*CsgB fibrils and the central repeats of *Ec*CsgA, too. Hence it is conceivable that the inner repeats of *Ec*CsgA R2, R3 and R4, which are less amyloidogenic, interact in an unspecific manner with *Ec*CsgB and hamper in that way their own correct fibril incorporation. However, the high intrinsic amyloidogenic propensity of R1 and R5 (Wang *et al.*, 2008b; Wang *et al.*, 2007) still enables these segments to fold regularly. This readily explains the observed structural pattern of the recently heteronucleated *Ec*CsgA$_{40\text{-}151}$ fibrils.

5.2.5 Aspects of Curli Biogenesis in CsgA Polymerisation

In vivo, the fibrillisation of the major curli subunit CsgA is highly regulated. The curli biogenesis is a sophisticated mechanism involving several accessory proteins illustrated in **chapter 1.3.1**, **Figure 1-10**. Since *in vivo* each accessory protein features an indispensible function with regard to the specific curli assembly (**Figure 5-12**), an *in vitro* effect on the CsgA fibril formation is imaginable, too.

After secretion to the extracellular space, CsgA is nucleated by CsgB into specific amyloid fibrils (Chapman *et al.*, 2002; Hammar *et al.*, 1996). Therefore, *in vitro* heteronucleated CsgA fibrils should be further investigated. Initial analysis of heteronucleated *Ec*CsgA$_{40\text{-}151}$ within this study revealed less homogeneity and less structural regularity of the fibrils, which was explained by the heterogeneity of *Ec*CsgB fibrils itself. So far, the *in vitro* assembly of *Ec*CsgB results always in heterogeneous fibrils (Madhu Nagaraj, unpublished). *In vivo*, CsgF assists the stability and therefore the regularity of CsgB that

nucleates CsgA (Nenninger *et al.*, 2009). Thus, at the cell surface, CsgA, CsgB and CsgF interact in a sophisticated manner with regard to the assembly of CsgA. Since CsgF is suggested to facilitate the entire CsgB nucleator activity, an *in vitro* effect is conceivable too. Therefore, it is obvious to investigate the effect of *Ec*CsgF on the homogeneity and regular structure of *Ec*CsgB and consequently the effect of these *Ec*CsgB fibrils on the nucleation and regular structure of *Ec*CsgA. Possibly, such an adjusted *in vitro* generation of *Ec*CsgA fibrils and variants facilitates further mutagenesis and fluorescence labelling studies of *Ec*CsgA to prove its suggested fold (**chapter 5.2.1**).

So far only less is known about the function of CsgC, but loss of CsgC results in aberrant curli fibril formation *in vivo* (Gibson *et al.*, 2007). Thus, an *in vitro* effect of CsgC on CsgA is arguable, too.

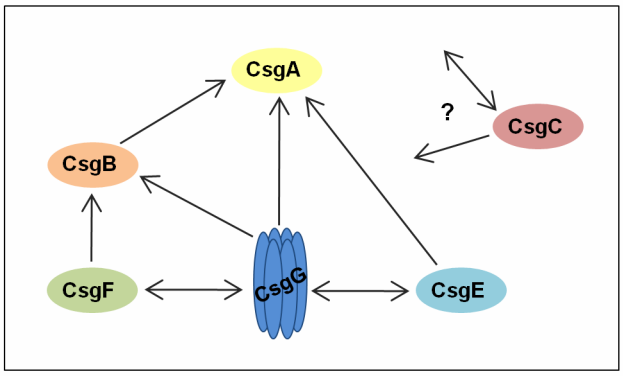

Figure 5-12: Accessory proteins during curli assembly and their interactions. CsgG forms an outer membrane channel and interacts with CsgE and CsgF. CsgE avoids premature CsgA assembly in the periplasm. CsgF is secreted via CsgG to the extracellular space, where it facilitates the cell association and stability of secreted CsgB. After secretion of CsgA, CsgB nucleates the polymerisation of CsgA into specific amyloid fibrils. For further explanations see **chapter 1.3.1** and citations therein.

5.2.6 Conclusion

In summary, the presented H/D exchange NMR and side-chain solvent accessibility analysis of *Ec*CsgA fibrils revealed the first sequence-specific information about structure elements within CsgA fibrils at all. The data enabled to propose a fold of *Ec*CsgA. Thus, *Ec*CsgA was indicated to form a β-solenoid structure with five windings per molecule and a strand-turn-strand motif per winding summarised in **Figure 5-13**. The data clearly demonstrated the repetitive structural character, which was already indicated by the characteristic repetitive sequence character (Collinson *et al.*, 1999). The intramolecularly conserved residues facilitate crucial functions providing the appropriate structure and the fold seems to be evolutionary optimised.

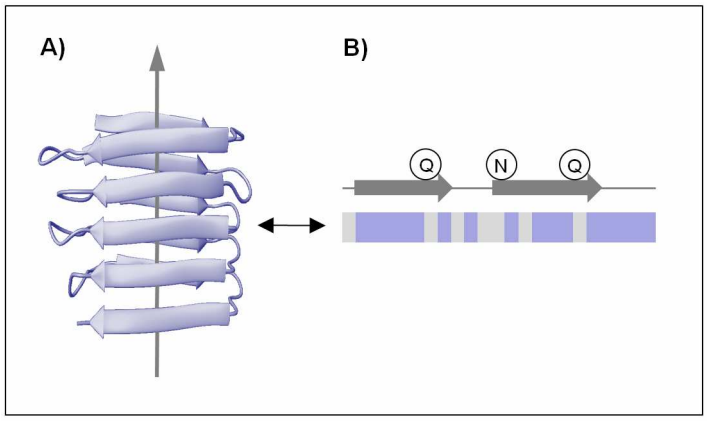

Figure 5-13: The proposed β-fold of *Ec*CsgA. (A) *Ec*CsgA was proposed to form a β-solenoid structure. The arrow indicates the fibril axis. **(B)** The repetitive sequence of *Ec*CsgA enables the β-solenoid structure. The amino acid sequence of the sequence repeats is shown as lilac bar. Intramolecularly conserved residues are shown in grey. Arrows above indicate identified β-strands. Probably stacking glutamine and asparagine residues are highlighted.

Since H/D exchange NMR spectroscopy is a sensitive tool to specify the homogeneity and stability of a structure, the analysis of two different *Ec*CsgA variants indicated biphasic and heterogeneous behaviour of distinct residues. This fact and the heterogeneity of heteronucleated fibrils were reconsidered and should be probed again with regard to conditions that are highly similar to the *in vivo* settings.

Similar to the H/D exchange NMR analysis of *Fg*HET-$s_{218-289}$ within this study, this technique was successfully implemented to gain residue-specific information about the structure, homogeneity and stability of the major curli subunit CsgA.

5.3 Monomeric Analogues of CsgA Oligomers – A Challenging Target

In order to obtain soluble amyloid-like intermediates of CsgA for structural analysis, probable monomeric analogues of *Ec*CsgA oligomers were designed and subjected to solubility and stability analyses. The proteins carried one or two copies of *Ec*CsgA lacking the N22 tag and the β-sheet aggregation was to be blocked by the negative design of soluble β-sheet rich proteins (Richardson and Richardson, 2002) as outlined in the experimental strategy in **chapter 4.1**. Altogether, 28 fusion proteins were tested to form soluble monomeric analogues of CsgA oligomers (**Figure 4-26**).

5.3.1 *Ec*CsgA Protein Design Resulted in Insoluble and Soluble Fusion Proteins

For proline and lysine variants of CsgA CM and CsgA CD fusion proteins no soluble protein was obtained at all. The proteins had to be purified under denaturing conditions and readily formed aggregates after buffer exchange to native conditions.

Contrary to the intent of generating soluble analogues of *Ec*CsgA oligomers, both proline and lysine variants of CsgA CM and CsgA CD fusion proteins seem to assemble like *Ec*CsgA wt into specific amyloid aggregates. In these cases, the negative design of natural β-sheet proteins was not employable. The introduction of lysine residues was expected to be readily applicable, because no backbone distortion has to be designed (Richardson and Richardson, 2002). In former studies, *de novo* amyloid fibrils already were converted into monomeric β-sheet proteins (Wang and Hecht, 2002). However, the introduction of lysine and proline residues to *Ec*CsgA fusion proteins resulted in amyloid-like aggregates indicating the high robustness of the sophisticated *Ec*CsgA fold. It is imaginable, that the designated cap strands are not well ordered (Cherny *et al.*, 2005) and loop out enabling intermolecular contacts of the inner segments (**Figure 5-14**).

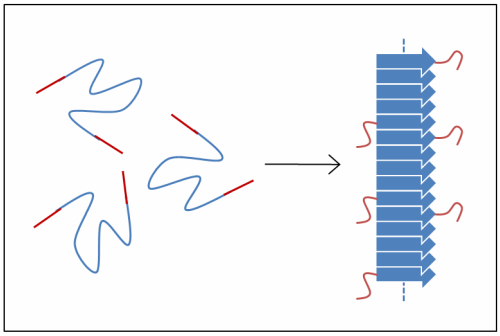

Figure 5-14: β-sheet in-corporation of designated cap strands. Native segments are shown in blue, red indicates designated cap segments. Arrows indicate β-sheet elements of each CsgA repeat. Sinuous lines present flexible segments.

Similar to CsgA CM and CsgA CD fusion proteins, the recombinant production of Bar-$(CsgA)_2$-Bar and Sumo-$(CsgA)_2$-Sumo resulted in highly stable aggregates indicating an

amyloid like conformation. In contrast to these insoluble expression products, Bar-CsgA-Bar and Sumo-CsgA-Sumo were partly soluble. Thus, these fusion proteins successfully altered the *Ec*CsgA amyloidogenic behaviour. For further discussion see **chapter 5.3.2**.

5.3.2 The Single *Ec*CsgA Moiety Is not Capable to Fold Regularly

The barnase and sumo fusion proteins of *Ec*CsgA, which carried only one copy of the CsgA moiety, were capable to inhibit fibril formation of *Ec*CsgA. Thus, the fused proteins blocked the intermolecular interactions of *Ec*CsgA segments that are crucial for appropriate self-polymerisation. But the recent experiments (illustrated in **chapter 4.4.4**) revealed that the fusion proteins are only soluble and folded to a certain extent. While the fusion partners are assumed to fold well similar to former studies on Ure2p fusion proteins (Baxa *et al.*, 2002), the single CsgA moiety is highly suggested to stay unfolded within the fusion proteins as proposed in **Figure 5-15** for Bar-CsgA-Bar. According to this, CsgA monomers are not capable to fold regularly. Former *in vitro* studies of CsgA self-polymerisation indicated the CsgA dimer to serve as fibrillisation nucleus (Wang *et al.*, 2007). Thus, CsgA monomers and dimers obviously differ in their structures and an unfolded monomer is conceivable. In consequence, CsgA oligomerisation involves CsgA folding. Blocking intermolecular contacts inhibits the folding event as shown for the barnase and sumo fusion proteins that carried only one copy of the CsgA moiety.

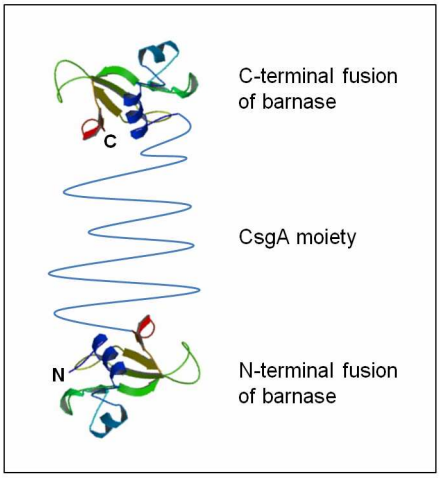

Figure 5-15: Schematic presentation of Bar-CsgA-Bar. Barnase is shown as ribbon diagram (pdb entry 1BNI, Buckle *et al.*, 1993). The probably unfolded CsgA moiety is shown as sinuous line.

In order to analyse the relationship between folding and oligomerisation, fusion proteins, which mimic the CsgA dimer, were also addressed to analyse. The corresponding fusion proteins Bar-$(CsgA)_2$-Bar and Sumo-$(CsgA)_2$-Sumo carried two copies of the CsgA moiety. These proteins always accumulated in highly stable inclusion bodies, which had to be solubilised in 90 % (v/v) formic acid indicating already amyloid character of the inclusion bodies. Inclusion bodies with an amyloid specific structure have been also shown for other proteins (Wang *et al.*, 2008a; Wasmer *et al.*, 2009a). Assuming an amyloid specific structure

of the inclusion bodies, which were formed by Bar-(CsgA)$_2$-Bar and Sumo-(CsgA)$_2$-Sumo, the mimicked CsgA dimer is capable to fold into the cross-β structure. Probably this results in a rearrangement of the fusion partners, barnase and sumo, enabling intermolecular contacts of the CsgA moieties and thus fibril formation. Another hypothesis is that the intracellular CsgA folding and fibrillisation event is much faster than folding of barnase and sumo disabling the designated blocking function of the fusion partners.

5.3.3 Conclusion

The effort to generate soluble analogues of *Ec*CsgA oligomers for structural analysis within this study revealed that it is quite challenging to obtain soluble and at the same time well-folded *Ec*CsgA analogues. The data clearly indicated that monomeric CsgA is not capable to fold regularly.

In order to obtain soluble and folded analogues of *Ec*CsgA oligomers distinct factors have to be considered. Since the strategy of blocking CsgA aggregation by covering the edge strands with well folding protein domains seems to be most promising, modifications of the available fusion proteins are conceivable. Gatekeeper residues control the folding behaviour of CsgA (Wang *et al.*, 2010). Mutations of these residues might increase the folding propensity of CsgA within the fusion proteins that carry one copy of the CsgA moiety. Furthermore, the fusion partners require a certain arrangement around the CsgA moiety to enable a successful capping. Therefore, the modification of the length of the corresponding linker peptides might be helpful to facilitate the required orientation of the capping proteins/domains.

6 Outlook

In the recent study, two amyloids were structurally investigated by H/D exchange NMR spectroscopy. The H/D exchange NMR analysis of *Fg*HET-s$_{218-289}$ and *Ec*CsgA enabled further insights into their fibril properties. Similar to the analyses of *Pa*HET-s$_{218-289}$ (Ritter *et al.*, 2005), Aβ_{1-42} (Luhrs *et al.*, 2005) and α-synuclein (Vilar *et al.*, 2008), the recent experiments revealed information about the sequence-specific positions of regular secondary structural elements, the stability and the homogeneity of the studied fibrils. Furthermore, in conjunction with complementary data, the almost complete H/D exchange rates enabled the suggestion of a probable fold of each studied protein. The H/D exchange NMR analyses within this study demonstrated differences in the homogeneity and stability of the investigated proteins. In case of heterogeneity, the determination of the H/D exchange rates still enabled the characterisation of the distinct structural entities, while heterogeneity would hamper other techniques. In summary, the H/D exchange NMR technique was successfully implemented to answer different questions within this study. Thus, H/D exchange NMR spectroscopy is a valuable tool for the analysis of amyloid structures and their stability. Especially in comparative analyses of homologous proteins, H/D exchange NMR spectroscopy readily provides meaningful information about structural similarities and differences, homogeneity and stability with regard to the protein settings.

The structure-function analysis of *Fg*HET-s$_{218-289}$ in comparison to its distinct homologue *Pa*HET-s$_{218-289}$ revealed that both proteins are structurally highly related. Similar to globular proteins (Chothia and Lesk, 1986; Ginalski, 2006), conserved residues enable the formation of a certain fold. So far, amyloids have been shown to form different polymorphic forms and much different structures already due to single point mutations (Tycko *et al.*, 2009). The present study strengthens the formation of highly evolved structures of functional amyloids, while disease related amyloids are often heterogeneous or polymorph (Greenwald and Riek, 2010; Ritter *et al.*, 2005). To prove this general aspect of functional and toxic amyloids, further structural analyses of distinct homologues of amyloids and prions should be followed. As illustrated in this study, fingerprint analyses via H/D exchange NMR spectroscopy are capable to identify appropriate similarities and differences with regard to structure, stability and homogeneity.

The H/D exchange NMR analysis of *Ec*CsgA within this study revealed the first sequence-specific information about structure elements of the major curli subunit at all.
H/D exchange NMR and side-chain solvent accessibility experiments revealed a possible fold of *Ec*CsgA within this study. Certainly, a complete solid state NMR analysis would answer remaining questions about the microstructure of CsgA. Furthermore, the H/D exchange NMR analysis of *Ec*CsgA fibrils revealed biphasic exchange behaviour for a couple of residues indicating local heterogeneity and different structural entities. In order to understand this behaviour, *Ec*CsgA fibrils should be also probed in the context of fragmentation and secondary nucleation at the surface of formed fibrils.
Since heteronucleated *Ec*CsgA fibrils revealed a limited regularity, homogeneity and stability in comparison to self-seeded fibrils, *in vitro* formed *Ec*CsgB fibrils were identified as no adequate seed. Therefore, seeding experiments with homogenous *Ec*CsgB fibrils in different concentrations would be helpful to analyse the effect of *Ec*CsgB fibrils on the structure and stability of *Ec*CsgA.
Curli fibrils bind to numerous host proteins, they facilitate manifold functions as potential pathogenicity factor of *Enterobacteria* (illustrated in **chapter 1.3.2** and reviewed by Barnhart and Chapman, 2006). Due to interactions with extracellular matrix proteins, curli support the attachment and invasion of host cells (Collinson *et al.*, 1993; Olsén *et al.*, 1993; Olsén *et al.*, 1989; Olsén *et al.*, 1998). Furthermore, curli bind to intracellular proteins (Bian *et al.*, 2000; Sjobring *et al.*, 1994) and contact-phase proteins (Ben Nasr *et al.*, 1996; Herwald *et al.*, 1998; Olsén *et al.*, 2002) enhancing the bacterial spreading throughout the host. As PAMPs, they are capable to induce the innate immune system (Froy, 2005; Kai-Larsen *et al.*, 2010; Tükel *et al.*, 2010). Since CsgA is the major curli subunit (Hammar *et al.*, 1995), CsgA mediated interactions are most likely. Obviously, interaction studies of CsgA with appropriate host proteins would provide further insights into the pathogenic mechanisms of curliated bacteria. The recent results provide the background for following structure-function analyses as they enable the precise mapping of binding positions to structural elements of CsgA.

Monomeric stable folded analogues of CsgA oligomers were identified as challenging target. Potentially, mutagenesis studies including the modification of gatekeeper residues within the CsgA moiety and a redesign of the CsgA fusion partners and the linker peptides of CsgA fusion proteins would ensure soluble and entire folded CsgA fusion proteins.

References

Aguzzi, A., Sigurdson, C., and Heikenwaelder, M. (2008). Molecular mechanisms of prion pathogenesis. Annu Rev Pathol *3*, 11-40.

Allen, K.D., Wegrzyn, R.D., Chernova, T.A., Muller, S., Newnam, G.P., Winslett, P.A., Wittich, K.B., Wilkinson, K.D., and Chernoff, Y.O. (2005). Hsp70 chaperones as modulators of prion life cycle: novel effects of Ssa and Ssb on the *Saccharomyces cerevisiae* prion [PSI+]. Genetics *169*, 1227-1242.

Andrew, E. (1981). Magic angle spinning. International Reviews in Physical Chemistry *1*, 195-224.

Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. Science *181*, 223-230.

Antzutkin, O.N., Balbach, J.J., Leapman, R.D., Rizzo, N.W., Reed, J., and Tycko, R. (2000). Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of beta-sheets in Alzheimer's beta-amyloid fibrils. Proc Natl Acad Sci U S A *97*, 13045-13050.

Arnold, U., and Ulbrich-Hofmann, R. (1999). Quantitative protein precipitation from guanidine hydrochloride-containing solutions by sodium deoxycholate/trichloroacetic acid. Anal Biochem *271*, 197-199.

Arnqvist, A., Olsen, A., Pfeifer, J., Russell, D.G., and Normark, S. (1992). The Crl protein activates cryptic genes for curli formation and fibronectin binding in *Escherichia coli* HB101. Mol Microbiol *6*, 2443-2452.

Assadi-Porter, F.M., Abildgaard, F., Blad, H., and Markley, J.L. (2003). Correlation of the Sweetness of Variants of the Protein Brazzein with Patterns of Hydrogen Bonds Detected by NMR Spectroscopy. Journal of Biological Chemistry *278*, 31331-31339.

Austin, J.W., Sanders, G., Kay, W.W., and Collinson, S.K. (1998). Thin aggregative fimbriae enhance *Salmonella enteritidis* biofilm formation. FEMS Microbiol Lett *162*, 295-301.

Ausubel, F.M., Brent, R., Kingston, R.E., Moor, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (2007). Current protocols in molecular biology. John Wiley & Sons Inc.

Bai, Y., Milne, J.S., Mayne, L., and Englander, S.W. (1993). Primary structure effects on peptide group hydrogen exchange. Proteins *17*, 75-86.

Balbach, J.J., Petkova, A.T., Oyler, N.A., Antzutkin, O.N., Gordon, D.J., Meredith, S.C., and Tycko, R. (2002). Supramolecular structure in full-length Alzheimer's beta-amyloid fibrils: evidence for a parallel beta-sheet organization from solid-state nuclear magnetic resonance. Biophys J *83*, 1205-1216.

Balguerie, A., Dos Reis, S., Ritter, C., Chaignepain, S., Coulary-Salin, B., Forge, V., Bathany, K., Lascu, I., Schmitter, J.M., Riek, R., *et al.* (2003). Domain organization and structure-function relationship of the HET-s prion protein of *Podospora anserina*. Embo J *22*, 2071-2081.

Barnhart, M.M., and Chapman, M.R. (2006). Curli biogenesis and function. Annu Rev Microbiol *60*, 131-147.

Baskakov, I.V., Legname, G., Baldwin, M.A., Prusiner, S.B., and Cohen, F.E. (2002). Pathway complexity of prion protein assembly into amyloid. J Biol Chem *277*, 21140-21148.

Baudet, S., and Janin, J. (1991). Crystal structure of a Barnase-d(GpC) complex at 1.9 Å resolution. J Mol Biol *219*, 123-132.

Bax, A., Griffey, R.H., and Hawkins, B.L. (1983). Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR* 1. Journal of Magnetic Resonance (1969) *55*, 301-315.

Bax, A., and Ikura, M. (1991). An efficient 3D NMR technique for correlating the proton and ^{15}N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins. J Biomol NMR *1*, 99-104.

Baxa, U., Speransky, V., Steven, A.C., and Wickner, R.B. (2002). Mechanism of inactivation on prion conversion of the Saccharomyces cerevisiae Ure2 protein. Proceedings of the National Academy of Sciences *99*, 5253-5260.

Baxa, U., Wickner, R.B., Steven, A.C., Anderson, D.E., Marekov, L.N., Yau, W.M., and Tycko, R. (2007). Characterization of beta-sheet structure in Ure2p$_{1-89}$ yeast prion fibrils by solid-state nuclear magnetic resonance. Biochemistry *46*, 13149-13162.

Ben Nasr, A., Olsen, A., Sjobring, U., Muller-Esterl, W., and Bjorck, L. (1996). Assembly of human contact phase proteins and release of bradykinin at the surface of curli-expressing *Escherichia coli*. Mol Microbiol *20*, 927-935.

Bensadoun, A., and Weinstein, D. (1976). Assay of proteins in the presence of interfering materials. Anal Biochem *70*, 241-250.

Berson, J.F., Harper, D.C., Tenza, D., Raposo, G., and Marks, M.S. (2001). Pmel17 initiates premelanosome morphogenesis within multivesicular bodies. Mol Biol Cell *12*, 3451-3464.

Berthomieu, C., and Hienerwadel, R. (2009). Fourier transform infrared (FTIR) spectroscopy. Photosynthesis research *101*, 157-170.

Bian, Z., Brauner, A., Li, Y., and Normark, S. (2000). Expression of and cytokine activation by *Escherichia coli* curli fibers in human sepsis. J Infect Dis *181*, 602-612.

Boehr, D.D., and Wright, P.E. (2008). How Do Proteins Interact? Science *320*, 1429.

Brok, R., Van Gelder, P., Winterhalter, M., Ziese, U., Koster, A.J., de Cock, H., Koster, M., Tommassen, J., and Bitter, W. (1999). The C-terminal domain of the *Pseudomonas* secretin XcpQ forms oligomeric rings with pore activity. J Mol Biol *294*, 1169-1179.

Brown, P., Liberski, P.P., Wolff, A., and Gajdusek, D.C. (1990). Conservation of infectivity in purified fibrillary extracts of scrapie-infected hamster brain after sequential enzymatic digestion or polyacrylamide gel electrophoresis. Proceedings of the National Academy of Sciences *87*, 7240-7244.

Buckle, A.M., Henrick, K., and Fersht, A.R. (1993). Crystal structural analysis of mutations in the hydrophobic cores of barnase. J Mol Biol *234*, 847-860.

Carulla, N., Caddy, G.L., Hall, D.R., Zurdo, J., Gairí, M., Feliz, M., Giralt, E., Robinson, C.V., and Dobson, C.M. (2005). Molecular recycling within amyloid fibrils. Nature *436*, 554-558.

Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., and Oschkinat, H. (2002). Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. Nature *420*, 98-102.

Chapman, M.R., Robinson, L.S., Pinkner, J.S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S.J. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. Science *295*, 851-855.

Chen, M., Margittai, M., Chen, J., and Langen, R. (2007). Investigation of alpha-synuclein fibril structure by site-directed spin labeling. J Biol Chem *282*, 24970-24979.

Cherny, I., Rockah, L., Levy-Nissenbaum, O., Gophna, U., Ron, E.Z., and Gazit, E. (2005). The formation of *Escherichia coli* curli amyloid fibrils is mediated by prion-like peptide repeats. J Mol Biol *352*, 245-252.

Chiba, T., Hagihara, Y., Higurashi, T., Hasegawa, K., Naiki, H., and Goto, Y. (2003). Amyloid fibril formation in the context of full-length protein: effects of proline mutations on the amyloid fibril formation of beta2-microglobulin. J Biol Chem *278*, 47016-47024.

Chien, P., Weissman, J.S., and DePace, A.H. (2004). Emerging principles of conformation-based prion inheritance. Annu Rev Biochem *73*, 617-656.

Chiti, F., and Dobson, C.M. (2006). Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem *75*, 333-366.

Chothia, C., and Lesk, A.M. (1986). The relation between the divergence of sequence and structure in proteins. Embo J *5*, 823-826.

Cohen, F.E., and Prusiner, S.B. (1998). Pathologic conformations of prion proteins. Annu Rev Biochem *67*, 793-819.

Coligan, J.E. (2003). Short Protocols in protein science. A compendium of methods from Current protocols in protein science. John Wiley & Sons Inc.

Collinson, S.K., Clouthier, S.C., Doran, J.L., Banser, P.A., and Kay, W.W. (1996). *Salmonella enteritidis* agfBAC operon encoding thin, aggregative fimbriae. J Bacteriol *178*, 662-667.

Collinson, S.K., Doig, P.C., Doran, J.L., Clouthier, S., Trust, T.J., and Kay, W.W. (1993). Thin, aggregative fimbriae mediate binding of *Salmonella enteritidis* to fibronectin. J Bacteriol *175*, 12-18.

Collinson, S.K., Parker, J.M., Hodges, R.S., and Kay, W.W. (1999). Structural predictions of AgfA, the insoluble fimbrial subunit of *Salmonella* thin aggregative fimbriae. J Mol Biol *290*, 741-756.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res *16*, 10881-10890.

Coustou, V., Deleu, C., Saupe, S., and Begueret, J. (1997). The protein product of the het-s heterokaryon incompatibility gene of the fungus *Podospora anserina* behaves as a prion analog. Proc Natl Acad Sci U S A *94*, 9773-9778.

Damaschun, G., Damaschun, H., Fabian, H., Gast, K., Krober, R., Wieske, M., and Zirwer, D. (2000). Conversion of yeast phosphoglycerate kinase into amyloid-like structure. Proteins *39*, 204-211.

Debnath, S., Chatterjee, S., Arif, M., Kundu, T.K., and Roy, S. (2011). Peptide-protein interactions suggest that acetylation of lysines 381 and 382 of p53 is important foa positive coactivator 4/p53 interaction. Journal of Biological Chemistry *286*, 25076-25087.

Deleu, C., Clave, C., and Begueret, J. (1993). A single amino acid difference is sufficient to elicit vegetative incompatibility in the fungus *Podospora anserina*. Genetics *135*, 45-52.

DePace, A.H., Santoso, A., Hillner, P., and Weissman, J.S. (1998). A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. Cell *93*, 1241-1252.

Der-Sarkissian, A., Jao, C.C., Chen, J., and Langen, R. (2003). Structural organization of alpha-synuclein fibrils studied by site-directed spin labeling. J Biol Chem *278*, 37530-37535.

Diercks, T., Coles, M., and Kessler, H. (1999). An efficient strategy for cross-peaks in 3D heteronuclear NOESY experiments. J Biomol NMR *15*, 177-180.

Dobson, C.M. (2004). Principles of protein folding, misfolding and aggregation. Semin Cell Dev Biol *15*, 3-16.

Dos Reis, S., Coulary-Salin, B., Forge, V., Lascu, I., Begueret, J., and Saupe, S.J. (2002). The HET-s prion protein of the filamentous fungus *Podospora anserina* aggregates in vitro into amyloid-like fibrils. J Biol Chem *277*, 5703-5706.

Drechsler, N., Fröbel, J., Jahreis, G., Gopalswamy, M., Balbach, J., Bosse-Doenecke, E., and Rudolph, R. (2011). Binding specificity of the ectodomain of the parathyroid hormone receptor. Biophysical Chemistry *154*, 66-72.

Dueholm, M.S., Petersen, S.V., Sonderkaer, M., Larsen, P., Christiansen, G., Hein, K.L., Enghild, J.J., Nielsen, J.L., Nielsen, K.L., Nielsen, P.H., *et al.* (2010). Functional amyloid in *Pseudomonas*. Mol Microbiol *77*, 1009-1020.

Edman, P., and Begg, G. (1967). A protein sequenator. Eur J Biochem *1*, 80-91.

Englander, S., Mayne, L., Bai, Y., and Sosnick, T. (1997). Hydrogen exchange: The modern legacy of Linderstrøm Lang. Protein Science *6*, 1101-1109.

Englander, S.W. (2000). Protein folding intermediates and pathways studied by hydrogen exchange. Annu Rev Biophys Biomol Struct *29*, 213-238.

Fandrich, M., Fletcher, M.A., and Dobson, C.M. (2001). Amyloid fibrils from muscle myoglobin. Nature *410*, 165-166.

Fandrich, M., Meinhardt, J., and Grigorieff, N. (2009). Structural polymorphism of Alzheimer Abeta and other amyloid fibrils. Prion *3*, 89-93.

Fowler, D.M., Koulov, A.V., Balch, W.E., and Kelly, J.W. (2007). Functional amyloid - from bacteria to humans. Trends Biochem Sci *32*, 217-224.

Froy, O. (2005). Regulation of mammalian defensin expression by Toll-like receptor-dependent and independent signalling pathways. Cell Microbiol *7*, 1387-1397.

Gardiennet, C., Henry, B., Kuad, P., Spiess, B., and Tekely, P. (2005). Straightforward detection of the secondary ionisation of the phosphate group and pK determinations by high-resolution solid-state ^{31}P NMR. Chem Commun (Camb), 180-182.

Gebbink, M.F., Claessen, D., Bouma, B., Dijkhuizen, L., and Wosten, H.A. (2005). Amyloids - a functional coat for microorganisms. Nat Rev Microbiol *3*, 333-341.

Gerstel, U., and Romling, U. (2001). Oxygen tension and nutrient starvation are major signals that regulate *agf*D promoter activity and expression of the multicellular morphotype in *Salmonella typhimurium*. Environ Microbiol *3*, 638-648.

Gerstel, U., and Romling, U. (2003). The *csg*D promoter, a control unit for biofilm formation in *Salmonella typhimurium*. Res Microbiol *154*, 659-667.

Gibson, D.L., White, A.P., Rajotte, C.M., and Kay, W.W. (2007). AgfC and AgfE facilitate extracellular thin aggregative fimbriae synthesis in *Salmonella Enteritidis*. Microbiology *153*, 1131-1140.

Ginalski, K. (2006). Comparative modeling for protein structure prediction. Curr Opin Struct Biol *16*, 172-177.

Glass, N.L., and Dementhon, K. (2006). Non-self recognition and programmed cell death in filamentous fungi. Curr Opin Microbiol *9*, 553-558.

Glass, N.L., Jacobson, D.J., and Shiu, P.K. (2000). The genetics of hyphal fusion and vegetative incompatibility in filamentous ascomycete fungi. Annu Rev Genet *34*, 165-186.

Gokhale, K.C., Newnam, G.P., Sherman, M.Y., and Chernoff, Y.O. (2005). Modulation of prion-dependent polyglutamine aggregation and toxicity by chaperone proteins in the yeast model. J Biol Chem *280*, 22809-22818.

Golde, T.E., Dickson, D., and Hutton, M. (2006). Filling the gaps in the Aβ cascade hypothesis of Alzheimer's disease. Current Alzheimer Research *3*, 421-430.

Greenwald, J., Buhtz, C., Ritter, C., Kwiatkowski, W., Choe, S., Maddelein, M.-L., Ness, F., Cescau, S., Soragni, A., Leitz, D., *et al.* (2010). The Mechanism of Prion Inhibition by HET-S. Molecular Cell *38*, 889-899.

Greenwald, J., and Riek, R. (2010). Biology of amyloid: structure, function, and regulation. Structure *18*, 1244-1260.

Gregersen, N., Bross, P., Vang, S., and Christensen, J.H. (2006). Protein misfolding and human disease. Annu Rev Genomics Hum Genet *7*, 103-124.

Grzesiek, S., and Bax, A. (1992). Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. J Magn Reson *96*, 432-440.

Guntert, P., Dotsch, V., Wider, C., and Wüthrich, K. (1992). Processing of multi-dimensional NMR data with new software PROSA. J Biomol NMR *2*, 619-629.

Hall-Stoodley, L., Costerton, J.W., and Stoodley, P. (2004). Bacterial biofilms: from the Natural environment to infectious diseases. Nat Rev Micro *2*, 95-108.

Hammar, M., Arnqvist, A., Bian, Z., Olsen, A., and Normark, S. (1995). Expression of two csg operons is required for production of fibronectin- and congo red-binding curli polymers in *Escherichia coli* K-12. Mol Microbiol *18*, 661-670.

Hammar, M., Bian, Z., and Normark, S. (1996). Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. Proc Natl Acad Sci U S A *93*, 6562-6566.

Hammer, N.D., Schmidt, J.C., and Chapman, M.R. (2007). The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. Proc Natl Acad Sci U S A *104*, 12494-12499.

Hansen, R.E., and Winther, J.R. (2009). An introduction to methods for analyzing thiols and disulfides: Reactions, reagents, and practical considerations. Anal Biochem *394*, 147-158.

Hardy, J., and Selkoe, D.J. (2002). The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. Science *297*, 353-356.

Harper, J.D., Lieber, C.M., and Lansbury, P.T., Jr. (1997). Atomic force microscopic imaging of seeded fibril formation and fibril branching by the Alzheimer's disease amyloid-beta protein. Chem Biol *4*, 951-959.

Hartl, F.U., and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. Nature Structural & Molecular Biology *16*, 574-581.

Hawe, A., Sutter, M., and Jiskoot, W. (2008). Extrinsic fluorescent dyes as tools for protein characterization. Pharmaceutical research *25*, 1487-1499.

Heise, H. (2008). Solid-state NMR spectroscopy of amyloid proteins. Chembiochem *9*, 179-189.

Heise, H., Hoyer, W., Becker, S., Andronesi, O.C., Riedel, D., and Baldus, M. (2005). Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. Proc Natl Acad Sci U S A *102*, 15871-15876.

Henderson, R. (2004). Realizing the potential of electron cryo-microscopy. Quarterly reviews of biophysics *37*, 3-13.

Hennel, J., and Klinowski, J. (2005). Magic-angle spinning: a historical perspective. New Techniques in Solid-State NMR, 1-14.

Hennetin, J., Jullian, B., Steven, A.C., and Kajava, A.V. (2006). Standard conformations of beta-arches in beta-solenoid proteins. J Mol Biol *358*, 1094-1105.

Herczenik, E., and Gebbink, M.F.B.G. (2008). Molecular and cellular aspects of protein misfolding and disease. The FASEB Journal *22*, 2115-2133.

Herwald, H., Morgelin, M., Olsen, A., Rhen, M., Dahlback, B., Muller-Esterl, W., and Bjorck, L. (1998). Activation of the contact-phase system on bacterial surfaces - a clue to serious complications in infectious diseases. Nat Med *4*, 298-302.

Hirakura, Y., Azimov, R., Azimova, R., and Kagan, B.L. (2000). Polyglutamine-induced ion channels: a possible mechanism for the neurotoxicity of Huntington and other CAG repeat diseases. J Neurosci Res *60*, 490-494.

Hirakura, Y., Carreras, I., Sipe, J.D., and Kagan, B.L. (2002). Channel formation by serum amyloid A: a potential mechanism for amyloid pathogenesis and host defense. Amyloid *9*, 13-23.

Hirakura, Y., Lin, M.C., and Kagan, B.L. (1999). Alzheimer amyloid Aβ1-42 channels: effects of solvent, pH, and Congo Red. J Neurosci Res *57*, 458-466.

Hirota-Nakaoka, N., Hasegawa, K., Naiki, H., and Goto, Y. (2003). Dissolution of 2-Microglobulin Amyloid Fibrils by Dimethylsulfoxide. J Biochem *134*, 159.

Hoashi, T., Watabe, H., Muller, J., Yamaguchi, Y., Vieira, W.D., and Hearing, V.J. (2005). MART-1 is required for the function of the melanosomal matrix protein PMEL17/GP100 and the maturation of melanosomes. J Biol Chem *280*, 14006-14016.

Hofmann, H., Weininger, U., Löw, C., Golbik, R.P., Balbach, J., and Ulbrich-Hofmann, R. (2008). Fast Amide Proton Exchange Reveals Close Relation between Native-State Dynamics and Unfolding Kinetics. J Am Chem Soc *131*, 140-146.

Hoshino, M., Katou, H., Hagihara, Y., Hasegawa, K., Naiki, H., and Goto, Y. (2002). Mapping the core of the β_2-microglobulin amyloid fibril by H/D exchange. Nat Struct Biol *9*, 332-336.

Hvidt, A., and Linderstrøm-Lang, K. (1954). Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. Biochim Biophys Acta *14*, 574.

Idnurm, A., and Howlett, B.J. (2003). Analysis of loss of pathogenicity mutants reveals that repeat-induced point mutations can occur in the Dothideomycete *Leptosphaeria maculans*. Fungal Genet Biol *39*, 31-37.

Jaenicke, R. (1996). How do proteins acquire their three-dimensional structure and stability? Naturwissenschaften *83*, 544-554.

Javitch, J.A., Shi, L., and Liapakis, G. (2002). Use of the substituted cysteine accessibility method to study the structure and function of G protein-coupled receptors. Methods Enzymol *343*, 137-156.

Jimenez, J.L., Guijarro, J.I., Orlova, E., Zurdo, J., Dobson, C.M., Sunde, M., and Saibil, H.R. (1999). Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. Embo J *18*, 815-821.

Jimenez, J.L., Nettleton, E.J., Bouchard, M., Robinson, C.V., Dobson, C.M., and Saibil, H.R. (2002). The protofilament structure of insulin amyloid fibrils. Proc Natl Acad Sci U S A *99*, 9196-9201.

Kai-Larsen, Y., Luthje, P., Chromek, M., Peters, V., Wang, X., Holm, A., Kadas, L., Hedlund, K.O., Johansson, J., Chapman, M.R., *et al.* (2010). Uropathogenic *Escherichia coli* modulates immune responses and its curli fimbriae interact with the antimicrobial peptide LL-37. PLoS Pathog *6*, e1001010.

Kajava, A.V., Baxa, U., and Steven, A.C. (2010). Beta arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils. Faseb J *24*, 1311-1319.

Kajava, A.V., and Steven, A.C. (2006). Beta-rolls, beta-helices, and other beta-solenoid proteins. In Adv Protein Chem (Academic Press) *73*, 55-96.

Kayed, R., Head, E., Thompson, J.L., McIntire, T.M., Milton, S.C., Cotman, C.W., and Glabe, C.G. (2003). Common Structure of Soluble Amyloid Oligomers Implies Common Mechanism of Pathogenesis. Science *300*, 486-489.

Keller, R. (2004). The Computer Aided Resonance Assignment Tutorial. CANTINA Verlag, Goldau, Switzerland.

Kheterpal, I., Cook, K.D., and Wetzel, R. (2006). Hydrogen/deuterium exchange mass spectrometry analysis of protein aggregates. Methods Enzymol *413*, 140-166.

Kheterpal, I., Williams, A., Murphy, C., Bledsoe, B., and Wetzel, R. (2001). Structural features of the Aβ amyloid fibril elucidated by limited proteolysis. Biochemistry *40*, 11757-11767.

Kourie, J.I. (2002). Prion channel proteins and their role in vacuolation and neurodegenerative diseases. Eur Biophys J *31*, 409-416.

Kourie, J.I., Farrelly, P.V., and Henry, C.L. (2001). Channel activity of deamidated isoforms of prion protein fragment 106-126 in planar lipid bilayers. J Neurosci Res *66*, 214-220.

Kourie, J.I., and Henry, C.L. (2002). Ion channel formation and membrane-linked pathologies of misfolded hydrophobic proteins: the role of dangerous unchaperoned molecules. Clin Exp Pharmacol Physiol *29*, 741-753.

Krebs, M.R., Bromley, E.H., and Donald, A.M. (2005). The binding of thioflavin-T to amyloid fibrils: localisation and implications. Journal of structural biology *149*, 30-37.

Krishnan, R., and Lindquist, S.L. (2005). Structural insights into a yeast prion illuminate nucleation and strain diversity. Nature *435*, 765-772.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Larsen, P., Nielsen, J.L., Dueholm, M.S., Wetzel, R., Otzen, D., and Nielsen, P.H. (2007). Amyloid adhesins are abundant in natural biofilms. Environ Microbiol *9*, 3077-3090.

Lashuel, H.A. (2005). Membrane permeabilization: a common mechanism in protein-misfolding diseases. Sci Aging Knowledge Environ *2005*, pe28.

Lashuel, H.A., Hartley, D., Petre, B.M., Walz, T., and Lansbury, P.T., Jr. (2002a). Neurodegenerative disease: amyloid pores from pathogenic mutations. Nature *418*, 291.

Lashuel, H.A., Petre, B.M., Wall, J., Simon, M., Nowak, R.J., Walz, T., and Lansbury, P.T., Jr. (2002b). Alpha-synuclein, especially the Parkinson's disease-associated mutants, forms pore-like annular and tubular protofibrils. J Mol Biol *322*, 1089-1102.

Lehrer, S.S. (1995). Pyrene excimer fluorescence as a probe of protein conformational change. Subcell Biochem *24*, 115-132.

Lehrer, S.S. (1997). Intramolecular pyrene excimer fluorescence: a probe of proximity and protein conformational change. Methods Enzymol *278*, 286-295.

LeMaster, D.M., and Richards, F.M. (1988). NMR sequential assignment of *Escherichia coli* thioredoxin utilizing random fractional deuteriation. Biochemistry *27*, 142-150.

LeVine, H., 3rd (1999). Quantification of beta-sheet amyloid fibril structures with thioflavin T. Methods Enzymol *309*, 274-284.

Li, R., and Woodward, C. (1999). The hydrogen exchange core and protein folding. Protein Sci *8*, 1571-1590.

Linderstrøm-Lang, K.U. (1952). Lane Medical Lectures: Proteins and Enzymes (Stanford University Press).

Linser, R., Dasari, M., Hiller, M., Higman, V., Fink, U., Lopez del Amo, J.-M., Markovic, S., Handel, L., Kessler, B., Schmieder, P., *et al.* (2011). Proton-Detected Solid-State NMR Spectroscopy of Fibrillar and Membrane Proteins. Angewandte Chemie International Edition *50*, 4508-4512.

Liwang, A.C., and Bax, A. (1997). Solution NMR Characterization of Hydrogen Bonds in a Protein by Indirect Measurement of Deuterium Quadrupole Couplings. J Magn Reson *127*, 54-64.

Loferer, H., Hammar, M., and Normark, S. (1997). Availability of the fibre subunit CsgA and the nucleator protein CsgB during assembly of fibronectin-binding curli is limited by the intracellular concentration of the novel lipoprotein CsgG. Mol Microbiol *26*, 11-23.

Lomakin, A., Teplow, D.B., Kirschner, D.A., and Benedek, G.B. (1997). Kinetic theory of fibrillogenesis of amyloid β-protein. Proceedings of the National Academy of Sciences *94*, 7942-7947.

Luhrs, T., Ritter, C., Adrian, M., Riek-Loher, D., Bohrmann, B., Dobeli, H., Schubert, D., and Riek, R. (2005). 3D structure of Alzheimer's amyloid-β(1-42) fibrils. Proc Natl Acad Sci U S A *102*, 17342-17347.

Maddelein, M.L., Dos Reis, S., Duvezin-Caubet, S., Coulary-Salin, B., and Saupe, S.J. (2002). Amyloid aggregates of the HET-s prion protein are infectious. Proc Natl Acad Sci U S A *99*, 7402-7407.

Makin, O.S., and Serpell, L.C. (2005). Structures for amyloid fibrils. FEBS Journal *272*, 5950-5961.

Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A., and Bax, A. (1989). Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. J Am Chem Soc *111*, 1515-1517.

Marley, J., Lu, M., and Bracken, C. (2001). A method for efficient isotopic labeling of recombinant proteins. J Biomol NMR *20*, 71-75.

Meinhardt, J., Sachse, C., Hortschansky, P., Grigorieff, N., and Fandrich, M. (2009). Aβ(1-40) fibril polymorphism implies diverse interaction patterns in amyloid fibrils. J Mol Biol *386*, 869-877.

Meyer, R.K., Lustig, A., Oesch, B., Fatzer, R., Zurbriggen, A., and Vandevelde, M. (2000). A monomer-dimer equilibrium of a cellular prion protein (PrP^{C}) not observed with recombinant PrP. Journal of Biological Chemistry *275*, 38081.

Mirzabekov, T.A., Lin, M.C., and Kagan, B.L. (1996). Pore formation by the cytotoxic islet amyloid peptide amylin. J Biol Chem *271*, 1988-1992.

Mori, S., Abeygunawardana, C., Johnson, M., and Vanzijl, P. (1995). Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. Journal of Magnetic Resonance, Series B *108*, 94-98.

Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E., and Dahlquist, F.W. (1989). Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. Methods Enzymol *177*, 44-73.

Murzin, A.G. (2008). Biochemistry. Metamorphic proteins. Science (New York, NY) *320*, 1725.

Nazabal, A., and Schmitter, J.M. (2006). Hydrogen-deuterium exchange analyzed by matrix-assisted laser desorption-ionization mass spectrometry and the HET-s prion model. Methods Enzymol *413*, 167-181.

Nelson, R., Sawaya, M.R., Balbirnie, M., Madsen, A.O., Riekel, C., Grothe, R., and Eisenberg, D. (2005). Structure of the cross-beta spine of amyloid-like fibrils. Nature *435*, 773-778.

Nenninger, A.A., Robinson, L.S., Hammer, N.D., Epstein, E.A., Badtke, M.P., Hultgren, S.J., and Chapman, M.R. (2011). CsgE is a curli secretion specificity factor that prevents amyloid fibre aggregation. Mol Microbiol *81*, 486-499.

Nenninger, A.A., Robinson, L.S., and Hultgren, S.J. (2009). Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF. Proc Natl Acad Sci U S A *106*, 900-905.

Nichols, M.R., Moss, M.A., Reed, D.K., Cratic-McDaniel, S., Hoh, J.H., and Rosenberry, T.L. (2005). Amyloid-β Protofibrils Differ from Amyloid-β Aggregates Induced in Dilute Hexafluoroisopropanol in Stability and Morphology. Journal of Biological Chemistry *280*, 2471-2480.

Ohhashi, Y., Ito, K., Toyama, B.H., Weissman, J.S., and Tanaka, M. (2010). Differences in prion strain conformations result from non-native interactions in a nucleus. Nat Chem Biol *6*, 225-230.

Olsén, A., Arnqvist, A., Hammar, M., Sukupolvi, S., and Normark, S. (1993). The RpoS sigma factor relieves H-NS-mediated transcriptional repression of csgA, the subunit gene of fibronectin-binding curli in *Escherichia coli*. Mol Microbiol *7*, 523-536.

Olsén, A., Herwald, H., Wikstrom, M., Persson, K., Mattsson, E., and Bjorck, L. (2002). Identification of two protein-binding and functional regions of curli, a surface organelle and virulence determinant of *Escherichia coli*. J Biol Chem *277*, 34568-34572.

Olsén, A., Jonsson, A., and Normark, S. (1989). Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*. Nature *338*, 652-655.

Olsén, A., Wick, M.J., Morgelin, M., and Bjorck, L. (1998). Curli, fibrous surface proteins of *Escherichia coli*, interact with major histocompatibility complex class I molecules. Infect Immun *66*, 944-949.

Osherovich, L.Z., Cox, B.S., Tuite, M.F., and Weissman, J.S. (2004). Dissection and design of yeast prions. PLoS Biol *2*, E86.

Otzen, D., and Nielsen, P.H. (2008). We find them here, we find them there: functional bacterial amyloid. Cell Mol Life Sci *65*, 910-927.

Petkova, A.T., Ishii, Y., Balbach, J.J., Antzutkin, O.N., Leapman, R.D., Delaglio, F., and Tycko, R. (2002). A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. Proc Natl Acad Sci U S A *99*, 16742-16747.

Petkova, A.T., Yau, W.M., and Tycko, R. (2006). Experimental constraints on quaternary structure in Alzheimer's β-amyloid fibrils. Biochemistry *45*, 498-512.

Prusiner, S. (1998). Prions. Proc Natl Acad Sci U S A *95*, 13363 - 13383.

Quist, A., Doudevski, I., Lin, H., Azimova, R., Ng, D., Frangione, B., Kagan, B., Ghiso, J., and Lal, R. (2005). Amyloid ion channels: a common structural link for protein-misfolding disease. Proc Natl Acad Sci U S A *102*, 10427-10432.

Ranjbar, B., and Gill, P. (2009). Circular Dichroism Techniques: Biomolecular and Nanostructural Analyses A Review. Chemical Biology & Drug Design *74*, 101-120.

Richardson, J.S., and Richardson, D.C. (2002). Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. Proc Natl Acad Sci U S A *99*, 2754-2759.

Ritter, C., Maddelein, M.L., Siemer, A.B., Luhrs, T., Ernst, M., Meier, B.H., Saupe, S.J., and Riek, R. (2005). Correlation of structural elements and infectivity of the HET-s prion. Nature *435*, 844-848.

Rizet, G. (1952). Les phénomènes de barrages chez *Podospora anserina*. I. Analyse génétique des barrages entre souches S et s. Rev Cytol Biol Vég *13*, 51-92.

Robinson, L.S., Ashman, E.M., Hultgren, S.J., and Chapman, M.R. (2006). Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. Mol Microbiol *59*, 870-881.

Rochet, J.-C., and Lansbury, P.T. (2000). Amyloid fibrillogenesis: themes and variations. Curr Opin Struct Biol *10*, 60-68.

Romero, D., Aguilar, C., Losick, R., and Kolter, R. (2010). Amyloid fibers provide structural integrity to Bacillus subtilis biofilms. Proceedings of the National Academy of Sciences *107*, 2230-2234.

Romling, U., Bian, Z., Hammar, M., Sierralta, W.D., and Normark, S. (1998a). Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. J Bacteriol *180*, 722-731.

Romling, U., Sierralta, W.D., Eriksson, K., and Normark, S. (1998b). Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the *agf*D promoter. Mol Microbiol *28*, 249-264.

Ross, E.D., Minton, A., and Wickner, R.B. (2005). Prion domains: sequences, structures and interactions. Nat Cell Biol *7*, 1039-1044.

Rousseau, F., Serrano, L., and Schymkowitz, J.W. (2006). How evolutionary pressure against protein aggregation shaped chaperone specificity. J Mol Biol *355*, 1037-1047.

Rudolph, R., Böhm, G., Lilie, H., and Jaenicke, R. (1997). Folding proteins. Protein function: A practical approach, 57-99.

Sabate, R., Baxa, U., Benkemoun, L., Sanchez de Groot, N., Coulary-Salin, B., Maddelein, M.L., Malato, L., Ventura, S., Steven, A.C., and Saupe, S.J. (2007). Prion and non-prion amyloids of the HET-s prion forming domain. J Mol Biol *370*, 768-783.

Sachse, C., Fandrich, M., and Grigorieff, N. (2008). Paired beta-sheet structure of an Aβ(1-40) amyloid fibril revealed by electron microscopy. Proc Natl Acad Sci U S A *105*, 7462-7466.

Sambrook, J., and Russell, D.W. (2000). Molecular cloning - a laboratory manual. Cold Spring Harbor Laboratory Press.

Saupe, S.J. (2000). Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. Microbiol Mol Biol Rev *64*, 489-502.

Saupe, S.J. (2011). The [Het-s] prion of *Podospora anserina* and its role in heterokaryon incompatibility. Semin Cell Dev Biol *In Press, Corrected Proof*.

Sawaya, M.R., Sambashivan, S., Nelson, R., Ivanova, M.I., Sievers, S.A., Apostol, M.I., Thompson, M.J., Balbirnie, M., Wiltzius, J.J., McFarlane, H.T., *et al.* (2007). Atomic structures of amyloid cross-beta spines reveal varied steric zippers. Nature *447*, 453-457.

Schmidt, M., Sachse, C., Richter, W., Xu, C., Fandrich, M., and Grigorieff, N. (2009). Comparison of Alzheimer Aβ(1-40) and Aβ(1-42) amyloid fibrils reveals similar protofilament structures. Proc Natl Acad Sci U S A *106*, 19813-19818.

Schmoldt, H.-U., Wentzel, A., Becker, S., and Kolmar, H. (2005). A fusion protein system for the recombinant production of short disulfide bond rich cystine knot peptides using barnase as a purification handle. Protein Expression and Purification *39*, 82-89.

Schulenburg, C., Löw, C., Weininger, U., Mrestani-Klaus, C., Hofmann, H., Balbach, J., Ulbrich-Hofmann, R., and Arnold, U. (2009). The Folding Pathway of Onconase Is Directed by a Conserved Intermediate. Biochemistry *48*, 8449-8457.

Shammas, C., Donarski, J.A., and Ramesh, V. (2007). NMR structure of the peptidyl transferase RNA inhibitor antibiotic amicetin. Magnetic Resonance in Chemistry *45*, 133-141.

Sheng, W., and Liao, X. (2002). Solution structure of a yeast ubiquitin-like protein Smt3: The role of structurally less defined sequences in protein–protein recognitions. Protein Science *11*, 1482-1491.

Shewmaker, F., McGlinchey, R.P., Thurber, K.R., McPhie, P., Dyda, F., Tycko, R., and Wickner, R.B. (2009). The functional curli amyloid is not based on in-register parallel β-sheet structure. J Biol Chem *284*, 25065-25076.

Shewmaker, F., Wickner, R.B., and Tycko, R. (2006). Amyloid of the prion domain of Sup35p has an in-register parallel beta-sheet structure. Proc Natl Acad Sci U S A *103*, 19754-19759.

Siemer, A.B., Arnold, A.A., Ritter, C., Westfeld, T., Ernst, M., Riek, R., and Meier, B.H. (2006a). Observation of highly flexible residues in amyloid fibrils of the HET-s prion. J Am Chem Soc *128*, 13224-13228.

Siemer, A.B., Ritter, C., Ernst, M., Riek, R., and Meier, B.H. (2005). High-resolution solid-state NMR spectroscopy of the prion protein HET-s in its amyloid conformation. Angew Chem Int Ed Engl *44*, 2441-2444.

Siemer, A.B., Ritter, C., Steinmetz, M.O., Ernst, M., Riek, R., and Meier, B.H. (2006b). ^{13}C, ^{15}N resonance assignment of parts of the HET-s prion protein in its amyloid form. J Biomol NMR *34*, 75-87.

Sjobring, U., Pohl, G., and Olsen, A. (1994). Plasminogen, absorbed by Escherichia coli expressing curli or by Salmonella enteritidis expressing thin aggregative fimbriae, can be activated by simultaneously captured tissue-type plasminogen activator (t-PA). Mol Microbiol *14*, 443-452.

Soto, C., Sigurdsson, E.M., Morelli, L., Kumar, R.A., Castano, E.M., and Frangione, B. (1998). β-sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: implications for Alzheimer's therapy. Nat Med *4*, 822-826.

Stryer, L. (1996). Biochemie / Lubert Stryer. Aus dem Engl. übers. von Günther Stoll . . . - 4. Aufl. - Heidelberg ; Berlin ; Oxford : Spektrum, Akad. Verl.

Sunde, M., Serpell, L.C., Bartlam, M., Fraser, P.E., Pepys, M.B., and Blake, C.C.F. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction1. J Mol Biol *273*, 729-739.

Thanassi, D.G., and Hultgren, S.J. (2000). Assembly of complex organelles: pilus biogenesis in gram-negative bacteria as a model system. Methods *20*, 111-126.

Thanassi, D.G., Saulino, E.T., Lombardo, M.J., Roth, R., Heuser, J., and Hultgren, S.J. (1998). The PapC usher forms an oligomeric channel: implications for pilus biogenesis across the outer membrane. Proc Natl Acad Sci U S A *95*, 3146-3151.

Thompson, L.D.R., Derringer, G.A., and Wenig, B.M. (2000). Amyloidosis of the Larynx: A Clinicopathologic Study of 11 Cases. Mod Pathol *13*, 528-535.

Torok, M., Milton, S., Kayed, R., Wu, P., McIntire, T., Glabe, C.G., and Langen, R. (2002). Structural and dynamic features of Alzheimer's Aβ peptide in amyloid fibrils studied by site-directed spin labeling. J Biol Chem *277*, 40810-40815.

Toyama, B.H., Kelly, M.J.S., Gross, J.D., and Weissman, J.S. (2007). The structural basis of yeast prion strain variants. Nature *449*, 233-237.

Toyama, B.H., and Weissman, J.S. (2011). Amyloid Structure: Conformational Diversity and Consequences. Annu Rev Biochem *80*, 557-585.

Tuite, M.F., and Cox, B.S. (2003). Propagation of yeast prions. Nat Rev Mol Cell Biol *4*, 878-890.

Tükel, Ç., Nishimori, J.H., Wilson, R.P., Winter, M.G., Keestra, A.M., Van Putten, J.P.M., and Bäumler, A.J. (2010). Toll-like receptors 1 and 2 cooperatively mediate immune responses to curli, a common amyloid from enterobacterial biofilms. Cell Microbiol *12*, 1495-1505.

Turcq, B., Deleu, C., Denayrolles, M., and Begueret, J. (1991). Two allelic genes responsible for vegetative incompatibility in the fungus *Podospora anserina* are not essential for cell viability. Mol Gen Genet *228*, 265-269.

Tycko, R. (2000). Solid-state NMR as a probe of amyloid fibril structure. Curr Opin Chem Biol *4*, 500-506.

Tycko, R. (2006). Solid-state NMR as a probe of amyloid structure. Protein Pept Lett *13*, 229-234.

Tycko, R. (2011). Solid State NMR Studies of Amyloid Fibril Structure. Annual review of physical chemistry *62*.

Tycko, R., Sciarretta, K.L., Orgel, J.P., and Meredith, S.C. (2009). Evidence for novel β-sheet structures in Iowa mutant β-amyloid fibrils. Biochemistry *48*, 6072-6084.

Uhlich, G.A., Keen, J.E., and Elder, R.O. (2001). Mutations in the *csg*D promoter associated with variations in curli expression in certain strains of *Escherichia coli* O157:H7. Appl Environ Microbiol *67*, 2367-2370.

Uptain, S.M., and Lindquist, S. (2002). Prions as protein-based genetic elements. Annual Reviews in Microbiology *56*, 703-741.

Valentine, R.C., Shapiro, B.M., and Stadtman, E.R. (1968). Regulation of glutamine synthetase. XII. Electron microscopy of the enzyme from *Escherichia coli*. Biochemistry *7*, 2143-2152.

Van Houdt, R., and Michiels, C.W. (2005). Role of bacterial cell surface structures in *Escherichia coli* biofilm formation. Res Microbiol *156*, 626-633.

Van Melckebeke, H., Wasmer, C., Lange, A., Ab, E., Loquet, A., Bockmann, A., and Meier, B.H. (2010). Atomic-resolution three-dimensional structure of HET-s(218-289) amyloid fibrils by solid-state NMR spectroscopy. J Am Chem Soc *132*, 13765-13775.

Vidal, O., Longin, R., Prigent-Combaret, C., Dorel, C., Hooreman, M., and Lejeune, P. (1998). Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression. J Bacteriol *180*, 2442-2449.

Vilar, M., Chou, H.T., Luhrs, T., Maji, S.K., Riek-Loher, D., Verel, R., Manning, G., Stahlberg, H., and Riek, R. (2008). The fold of alpha-synuclein fibrils. Proc Natl Acad Sci U S A *105*, 8637-8642.

Volles, M.J., and Lansbury, P.T., Jr. (2002). Vesicle permeabilization by protofibrillar alpha-synuclein is sensitive to Parkinson's disease-linked mutations and occurs by a pore-like mechanism. Biochemistry *41*, 4595-4602.

Volles, M.J., Lee, S.J., Rochet, J.C., Shtilerman, M.D., Ding, T.T., Kessler, J.C., and Lansbury, P.T., Jr. (2001). Vesicle permeabilization by protofibrillar alpha-synuclein: implications for the pathogenesis and treatment of Parkinson's disease. Biochemistry *40*, 7812-7819.

Wang, L., Maji, S.K., Sawaya, M.R., Eisenberg, D., and Riek, R. (2008a). Bacterial inclusion bodies contain amyloid-like structure. PLoS Biol *6*, e195.

Wang, W., and Hecht, M.H. (2002). Rationally designed mutations convert de novo amyloid-like fibrils into monomeric β-sheet proteins. Proc Natl Acad Sci U S A *99*, 2760-2765.

Wang, X., Hammer, N.D., and Chapman, M.R. (2008b). The Molecular Basis of Functional Bacterial Amyloid Polymerization and Nucleation. Journal of Biological Chemistry *283*, 21530-21539.

Wang, X., Smith, D.R., Jones, J.W., and Chapman, M.R. (2007). In vitro polymerization of a functional *Escherichia coli* amyloid protein. J Biol Chem *282*, 3713-3719.

Wang, X., Zhou, Y., Ren, J.J., Hammer, N.D., and Chapman, M.R. (2010). Gatekeeper residues in the major curlin subunit modulate bacterial amyloid fiber biogenesis. Proc Natl Acad Sci U S A *107*, 163-168.

Wasmer, C., Benkemoun, L., Sabate, R., Steinmetz, M.O., Coulary-Salin, B., Wang, L., Riek, R., Saupe, S.J., and Meier, B.H. (2009a). Solid-state NMR spectroscopy reveals that E. coli inclusion bodies of HET-s(218-289) are amyloids. Angew Chem Int Ed Engl *48*, 4858-4860.

Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A.B., Riek, R., and Meier, B.H. (2008). Amyloid fibrils of the HET-s(218-289) prion form a β-solenoid with a triangular hydrophobic core. Science *319*, 1523-1526.

Wasmer, C., Schutz, A., Loquet, A., Buhtz, C., Greenwald, J., Riek, R., Bockmann, A., and Meier, B.H. (2009b). The molecular organization of the fungal prion HET-s in its amyloid form. J Mol Biol *394*, 119-127.

Wasmer, C., Zimmer, A., Sabate, R., Soragni, A., Saupe, S.J., Ritter, C., and Meier, B.H. (2010). Structural similarity between the prion domain of HET-s and a homologue can explain amyloid cross-seeding in spite of limited sequence identity. J Mol Biol *402*, 311-325.

Westermark, G.T., Johnson, K.H., and Westermark, P. (1999). Staining methods for identification of amyloid in tissue. In Methods Enzymol, W. Ronald, ed. (Academic Press), pp. 3-25.

Wickner, R.B., Dyda, F., and Tycko, R. (2008a). Amyloid of Rnq1p, the basis of the [PIN+] prion, has a parallel in-register β-sheet structure. Proc Natl Acad Sci U S A *105*, 2403-2408.

Wickner, R.B., Edskes, H.K., Shewmaker, F., Kryndushkin, D., and Nemecek, J. (2009). Prion variants, species barriers, generation and propagation. J Biol *8*, 47.

Wickner, R.B., Edskes, H.K., Shewmaker, F., and Nakayashiki, T. (2007). Prions of fungi: inherited structures and biological roles. Nature Reviews Microbiology *5*, 611-618.

Wickner, R.B., Shewmaker, F., Kryndushkin, D., and Edskes, H.K. (2008b). Protein inheritance (prions) based on parallel in-register β-sheet amyloid structures. Bioessays *30*, 955-964.

Williams, A.D., Portelius, E., Kheterpal, I., Guo, J.T., Cook, K.D., Xu, Y., and Wetzel, R. (2004). Mapping Aβ amyloid fibril secondary structure using scanning proline mutagenesis. J Mol Biol *335*, 833-842.

Wiltzius, J.J., Landau, M., Nelson, R., Sawaya, M.R., Apostol, M.I., Goldschmidt, L., Soriaga, A.B., Cascio, D., Rajashankar, K., and Eisenberg, D. (2009). Molecular mechanisms for protein-encoded inheritance. Nature Structural & Molecular Biology *16*, 973-978.

Wiltzius, J.J., Sievers, S.A., Sawaya, M.R., Cascio, D., Popov, D., Riekel, C., and Eisenberg, D. (2008). Atomic structure of the cross-beta spine of islet amyloid polypeptide (amylin). Protein Sci *17*, 1467-1474.

Wishart, D.S., and Nip, A.M. (1998). Protein chemical shift analysis: a practical guide. Biochem Cell Biol *76*, 153-163.

Wittekind, M., and Mueller, L. (1993). HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha-and beta-carbon resonances in proteins. Journal of magnetic resonance Series B *101*, 201-205.

Wood, S.J., Wetzel, R., Martin, J.D., and Hurle, M.R. (1995). Prolines and amyloidogenicity in fragments of the Alzheimer's peptide Aβ. Biochemistry *34*, 724-730.

Wormell, R.L. (1954). New fibres from proteins (Academic Press).

Wüthrich, K. (1990). Protein structure determination in solution by NMR spectroscopy. Journal of Biological Chemistry *265*, 22059-22062.

Yankner, B., Duffy, L., and Kirschner, D. (1990). Neurotrophic and neurotoxic effects of amyloid beta protein: reversal by tachykinin neuropeptides. Science *250*, 279-282.

Yoder, M., and Jurnak, F. (1995). Protein motifs. 3. The parallel beta helix and other coiled folds. The FASEB Journal *9*, 335-342.

Zahn, R., Buckle, A.M., Perrett, S., Johnson, C.M., Corrales, F.J., Golbik, R., and Fersht, A.R. (1996). Chaperone activity and structure of monomeric polypeptide binding domains of GroEL. Proc Natl Acad Sci U S A *93*, 15024-15029.

Zandomeneghi, G., Krebs, M.R., McCammon, M.G., and Fandrich, M. (2004). FTIR reveals structural differences between native beta-sheet proteins and amyloid fibrils. Protein Sci *13*, 3314-3321.

Zeeman, P. (1897). The effect of magnetisation on the nature of light emitted by a substance. Nature *55*, 347-347.

Zhang, R., Hu, X., Khant, H., Ludtke, S.J., Chiu, W., Schmid, M.F., Frieden, C., and Lee, J.M. (2009). Interprotofilament interactions between Alzheimer's Aβ1-42 peptides in amyloid fibrils revealed by cryoEM. Proc Natl Acad Sci U S A *106*, 4653-4658.

Zogaj, X., Bokranz, W., Nimtz, M., and Romling, U. (2003). Production of cellulose and curli fimbriae by members of the family *Enterobacteriaceae* isolated from the human gastrointestinal tract. Infect Immun *71*, 4151-4158.

Appendix

A) Amino acid sequences

CsgA (curli subunit) [*E. coli*] – amino acid sequence (GenBank entry AAA23616)

```
  1 MKLLKVAAIA AIVFSGSALA GVVPQYGGGG NHGGGGNNSG PNSELNIYQY GGGNSALALQ
 61 TDARNSDLTI TQHGGGNGAD VGQGSDDSSI DLTQRGFGNS ATLDQWNGKN SEMTVKQFGG
121 GNGAAVDQTA SNSSVNVTQV GFGNNATAHQ Y
```

CsgB (nucleation component of curli monomers) [*E. coli*]– amino acid sequence (GenBank entry CAA62281)

```
  1 MKNKLLFMML TILGAPGIAA AAGYDLANSE YNFAVNELSK SSFNQAAIIG QAGTNNSAQL
 61 RQGGSKLLAV VAQEGSSNRA KIDQTGDYNL AYIDQAGSAN DASISQGAYG NTAMIIQKGS
121 GNKANITQYG TQKTAIVVQR QSQMAIRVTQ R
```

Het-s (hypothetical protein) [Fusarium graminearum] – amino acid sequence (GenBank entry FG10600)

```
  1 MAEIFGIVSG ALSVAAIFNN CVDTFEYIQL GRRFGEDFQR YQLKLDLAKT RLGRWGEAIS
 61 INNEPRFSSF ASADKEVNIA REILEDIASC FEGAQKKSSR YADRADQGEL EIFGESDMNP
121 MLRRLHRHSK DIARQRQKTT SIIKKTKWAL YDAKSLERTI DQICSWIDEL EKLFPEQSAQ
181 TQLVEREIEK IDDKPTLEAL KDAASGVDPV MEDAVQRKLN MIEGHNSAEF VNLEGSAKFL
241 VGNVFSEKFL QRDVLLNDRT KNSMRTVSAT NQSRLQVGNV YGGRGIWED
```

Barnase (RNase) precursor [Bacillus amyloliquefaciens] – amino acid sequence (GenBank entry AAA86441.1)

```
  1 MKKRLSWISV CLLVLVSAAG MLFSTAAKTE TSSHKAHTEA QVINTFDGVA DYLQTYHKLP
 61 DNYITKSEAQ ALGWVASKGN LADVAPGKSI GGDIFSNREG KLPGKSGRTW READINYTSG
121 FRNSDRILYS SDWLIYKTTD HYQTFTKIR
```

Ubiquitin-like protein SMT3 [Saccharomyces cerevisiae] – amino acid sequence (GenBank entry AAB64951) (Sumo)

```
  1 MSDSEVNQEA KPEVKPEVKP ETHINLKVSD GSSEIFFKIK KTTPLRRLME AFAKRQGKEM
 61 DSLRFLYDGI RIQADQTPED LDMEDNDIIE AHREQIGGAT Y
```

Recombinant expression constructs

All constructs carried a hexa histidine tag, highlighted in blue.

$CsgA_{21\text{-}151}$ – amino acid sequence

```
 20 MGVVPQYGGG GNHGGGGNNS GPNSELNIYQ YGGGNSALAL QTDARNSDLT ITQHGGGNGA
 80 DVGQGSDDSS IDLTQRGFGN SATLDQWNGK NSEMTVKQFG GGNGAAVDQT ASNSSVNVTQ
140 VGFGNNATAH QYHHHHHH
```

$CsgA_{40\text{-}151}$ – amino acid sequence

```
 39 MGPNSELNIY QYGGGNSALA LQTDARNSDL TITQHGGGNG ADVGQGSDDS SIDLTQRGFG
 99 NSATLDQWNG KNSEMTVKQF GGGNGAAVDQ TASNSSVNVT QVGFGNNATA HQYHHHHHH
```

CsgA CM – amino acid sequence

```
  1 MRGSHHHHHH GLVPRGSNSG PNSSVNVTQV GFGNNATAHQ YSGDNSELNI YQYGGGNSAL
 61 ALQTDARNSD LTITQHGGGN GADVGQGSDD SSIDLTQRGF GNSATLDQWN GKNSEMTVKQ
121 FGGGNGAAVD QTASNSSVNV TQVGFGNNAT AHQYSGDNSE LNIYQYGGGN SALALQT
```

CsgA CD – amino acid sequence

```
  1 MRGSHHHHHH GLVPRGSNSG PNSSVNVTQV GFGNNATAHQ YSGDNSELNIY QYGGGNSALA
 61 LQTDARNSDL TITQHGGGNG ADVGQGSDDS SIDLTQRGFG NSATLDQWNGK NSEMTVKQFG
121 GGNGAAVDQT ASNSSVNVTQ VGFGNNATAH QYSGDNSELN IYQYGGGNSAL ALQTDARNSD
181 LTITQHGGGN GADVGQGSDD SSIDLTQRGF GNSATLDQWN GKNSEMTVKQ FGGGNGAAVD
241 QTASNSSVNV TQVGFGNNAT AHQYSGDNSE LNIYQYGGGN SALALQT
```

Bar-CsgA-Bar – amino acid sequence

```
  1 MGAQVINTFD GVADYLQTYH KLPDNYITKS EAQALGWVAS KGNLADVAPG KSIGGDIFSN
 61 REGKLPGKSG RTWREADINY TSGFRNSDRI LYSSDWLIYK TTDAYQTFTK IRGSGPNSEL
121 NIYQYGGGNS ALALQTDARN SDLTITQHGG GNGADVGQGS DDSSIDLTQR GFGNSATLDQ
181 WNGKNSEMTV KQFGGGNGAA VDQTASNSSV NVTQVGFGNN ATAHQYSSSM SAQVSNTFDG
241 VADYLQTYHK LPDNYITKSE AQALGWVASK GNLADVAPGK SIGGDIFSNR EGKLPGKSGR
301 TWREADINYT SGFRNSDRIL YSSDWLIYKT TDAYQTFTKI RHHHHHH
```

Bar-$(CsgA)_2$-Bar – amino acid sequence

```
  1 MGAQVINTFD GVADYLQTYH KLPDNYITKS EAQALGWVAS KGNLADVAPG KSIGGDIFSN
 61 REGKLPGKSG RTWREADINY TSGFRNSDRI LYSSDWLIYK TTDAYQTFTK IRGSGPNSEL
121 NIYQYGGGNS ALALQTDARN SDLTITQHGG GNGADVGQGS DDSSIDLTQR GFGNSATLDQ
181 WNGKNSEMTV KQFGGGNGAA VDQTASNSSV NVTQVGFGNN ATAHQYSGDN SELNIYQYGG
241 GNSALALQTD ARNSDLTITQ HGGGNGADVG QGSDDSSIDL TQRGFGNSAT LDQWNGKNSE
301 MTVKQFGGGN GAAVDQTASN SSVNVTQVGF GNNATAHQSS SMSAQVSNTF DGVADYLQTY
361 HKLPDNYITK SEAQALGWVA SKGNLADVAP GKSIGGDIFS NREGKLPGKS GRTWREADIN
421 YTSGFRNSDR ILYSSDWLIY KTTDAYQTFT KIRHHHHHH
```

Sumo-CsgA-Sumo – amino acid sequence

```
  1 MGHHHHHHGS LQDSEVNQEA KPEVKPEVKP ETHINLKVSD GSSEIFFKIK KTTPLRRLME
 61 AFAKRQGKEM DSLTFLYDGI EIQADQTPED LDMEDNDIIE AHREQIGGMP GRSGPNSELN
121 IYQYGGGNSA LALQTDARNS DLTITQHGGG NGADVGQGSD DSSIDLTQRG FGNSATLDQW
181 NGKNSEMTVK QFGGGNGAAV DQTASNSSVN VTQVGFGNNA TAHQYSSSMS GSLQDSEVNQ
241 EAKPEVKPEV KPETHINLKV SDGSSEIFFK IKKTTPLRRL MEAFAKRQGK EMDSLTFLYD
301 GIEIQADQTP EDLDMEDNDI IEAHREQIGG MPGR
```

Sumo-$(CsgA)_2$-Sumo – amino acid sequence

```
  1 MGHHHHHHGS LQDSEVNQEA KPEVKPEVKP ETHINLKVSD GSSEIFFKIK KTTPLRRLME
 61 AFAKRQGKEM DSLTFLYDGI EIQADQTPED LDMEDNDIIE AHREQIGGMP GRSGPNSELN
121 IYQYGGGNSA LALQTDARNS DLTITQHGGG NGADVGQGSD DSSIDLTQRG FGNSATLDQW
181 NGKNSEMTVK QFGGGNGAAV DQTASNSSVN VTQVGFGNNA TAHQYSGDNS ELNIYQYGGG
241 NSALALQTDA RNSDLTITQH GGGNGADVGQ GSDDSSIDLT QRGFGNSATL DQWNGKNSEM
301 TVKQFGGGNG AAVDQTASNS SVNVTQVGFG NNATAHQSSS MSGSLQDSEV NQEAKPEVKP
361 EVKPETHINL KVSDGSSEIF FKIKKTTPLR RLMEAFAKRQ GKEMDSLTFL YDGIEIQADQ
421 TPEDLDMEDN DIIEAHREQI GGMPGR
```

***Fg*Het-$s_{218\text{-}289}$ – amino acid sequence**

```
217 MKLNMIEGHN SAEFVNLEGS AKFLVGNVFS EKFLQRDVLL NDRTKNSMRT VSATNQSRLQ
277 VGNVYGGRGI WEDHHHHHH
```

B) Supplementary Methods

Table App-1: Conditions for labelling with thiol-reactive agents.

Native conditions (post assembly)					
No.	**Desalting**	**Fibrillisation**	**Labelling buffer**	**Labelling agent**	**Molar excess**
1	50 mM KPi pH 7.2	2.5 % (w/w) wt seed, mild agitation	50 mM KPi pH 7.2	AF488	2.5 x
				PM	3 x
2	50 mM Tris/HCl pH 7.5, 150 mM NaCl	2.5 % (w/v) wt seed, 5 mM TCEP, mild agitation	50 mM Tris/HCl pH 7.5, 150 mM NaCl	AF488	2.5 x
				PM	3 x
3	50 mM Tris/HCl pH 7.5, 150 mM NaCl, presence of 10 %(w/w) wt	2.5 % (w/v) wt seed, 5 mM TCEP, mild agitation	50 mM Tris/HCl pH 7.5, 150 mM NaCl	AF488	2.5 x
				PM	3 x
4	50 mM KPi pH 7.2, 5 mM DTT, presence of 10 %(w/w) wt	without seed, no agitation	50 mM KPi pH 7.2	PM	3 x
5	50 mM KPi pH 7.2, 5 mM DTT, presence of 10 %(w/w) wt	2.5 % (w/w) wt seed, mild agitation	50 mM Tris/HCl pH 8, 150 mM NaCl, 0.1 mM TCEP	AF488	2.5 x
				PM	10 x
				IAA	20 x
				IAM	20 x

Denaturing conditions (pre assembly)					
No.	**Labelling buffer**	**Labelling agent**	**Molar excess**	**Desalting**	**Fibrillisation**
6	50 mM KPi pH 7.2, 150 mM NaCl, 8 M GdmHCl	PM	3 x	20 mM DTT added to labelling sample, dialysis against 50 mM KPi pH 7.2	dialysis
7	50 mM Tris/HCl pH 8, 8 M GdmHCl 150 mM NaCl 5 mM TCEP	PM	3 x	20 mM DTT added to labelling sample, 50 mM Tris/HCl pH 7.5, 150 mM NaCl	2.5 mM (w/w) wt seed, mild agitation
8	50 mM Tris/HCl pH 6, 8 M GdmHCl 150 mM NaCl, 5 x DTT	IAA	25 x	10 mM CH_3COONa pH 7.5	-
		IAM	25 x		
		PM	15 x		
No.	**Labelling buffer**	**Labelling agent**	**Molar excess**	**Desalting**	**Fibrillisation**
9	50 mM Tris/HCl pH 6, 8 M GdmHCl 150 mM NaCl, 5 x TCEP	IAA	25 x	10 mM CH_3COONa pH 7.5	-
		IAM	25 x		
		PM	15 x		
10	50 mM Tris/HCl pH 6, 8 M GdmHCl 150 mM NaCl, 20 x DTT	IAA	100 x	10 mM CH_3COONa pH 7.5	-
		IAM	100 x		

Prior labelling studies of fibrils, protein samples always were desalted via Nap5 after reduction of the thiol groups as described in **chapter 3.4.5**. Under native conditions 20 - 100 µM of two days old fibrils were incubated in presence of at least 2.5 molar excess of an appropriate thiol-reactive agent for 30 min (AF488) and 3 h (IAA, IAM, PM) respectively. The sample volume was around 250 - 500 µl. The labelling reaction was quenched by centrifugation (20800 g, 5 min, RT) and three washing steps comprising repetitive re-suspension in 50 mM KPi pH 7.2 and centrifugation (20800 g, 5 min, RT). In the first washing step the buffer was supplemented with 5 mM DTT.
Under denaturing conditions 100 – 200 µM protein were applied to an appropriate thiol-reactive agent. The sample volume was around 100 – 500 µl. The labelling reaction was quenched by adding 20 mM DTT. In order to desalt the samples, they were dialysed or applied to a Nap5 column (**Table 3-13**).

C) Supplementary Figures

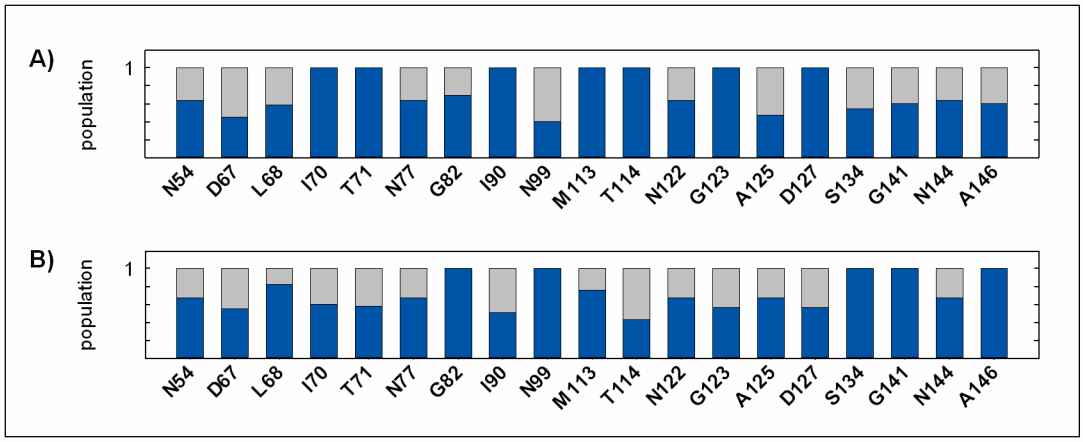

Figure App-6-1: Relative exchange behaviour and population of *Ec*CsgA residues with biphasic H/D exchange behaviour. Fast exchange is shown in grey, slow exchange in blue. **(A)** *Ec*CsgA$_{40\text{-}151}$ fibrils. **(B)** *Ec*CsgA$_{21\text{-}151}$ fibrils.

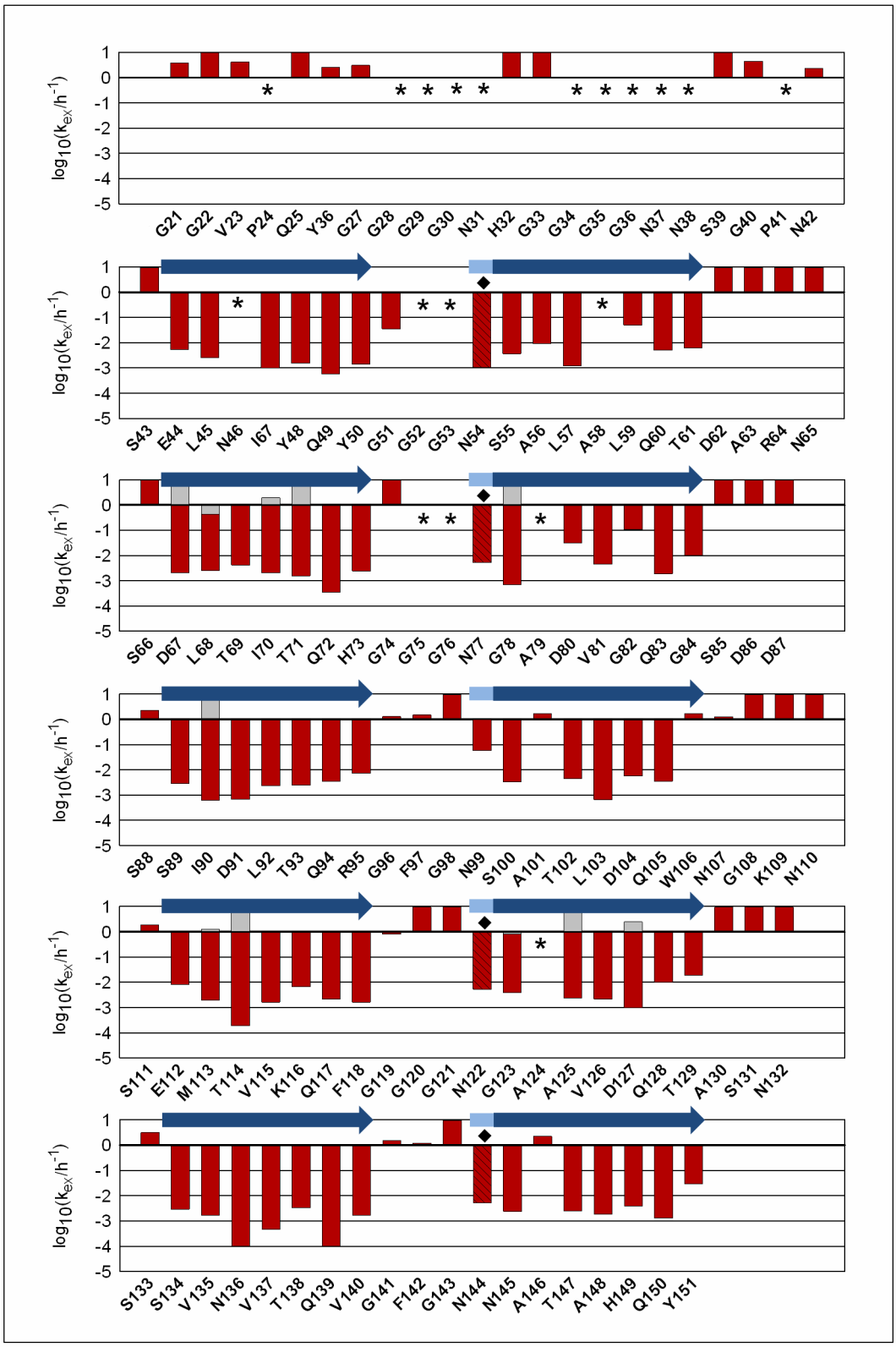

Figure App-6-2: H/D exchange NMR data of *Ec*CsgA$_{21\text{-}151}$ and secondary structure prediction. Red and grey bars indicate exchange rates of *Ec*CsgA$_{21\text{-}151}$ k_{ex} (h^{-1}). Grey bars indicate exchange rates of the fast first phase in case of biphasic behaviour. Patterned bars present the exchange rates of the asparagine ladder. Residues marked with a rectangle were analysed as one unit. Residues marked with an asterisk were excluded from the analysis due to fast intrinsic exchange in DMSO (N46, A58, A79, A124) and missing assignments. Blue arrows indicate β-strands.

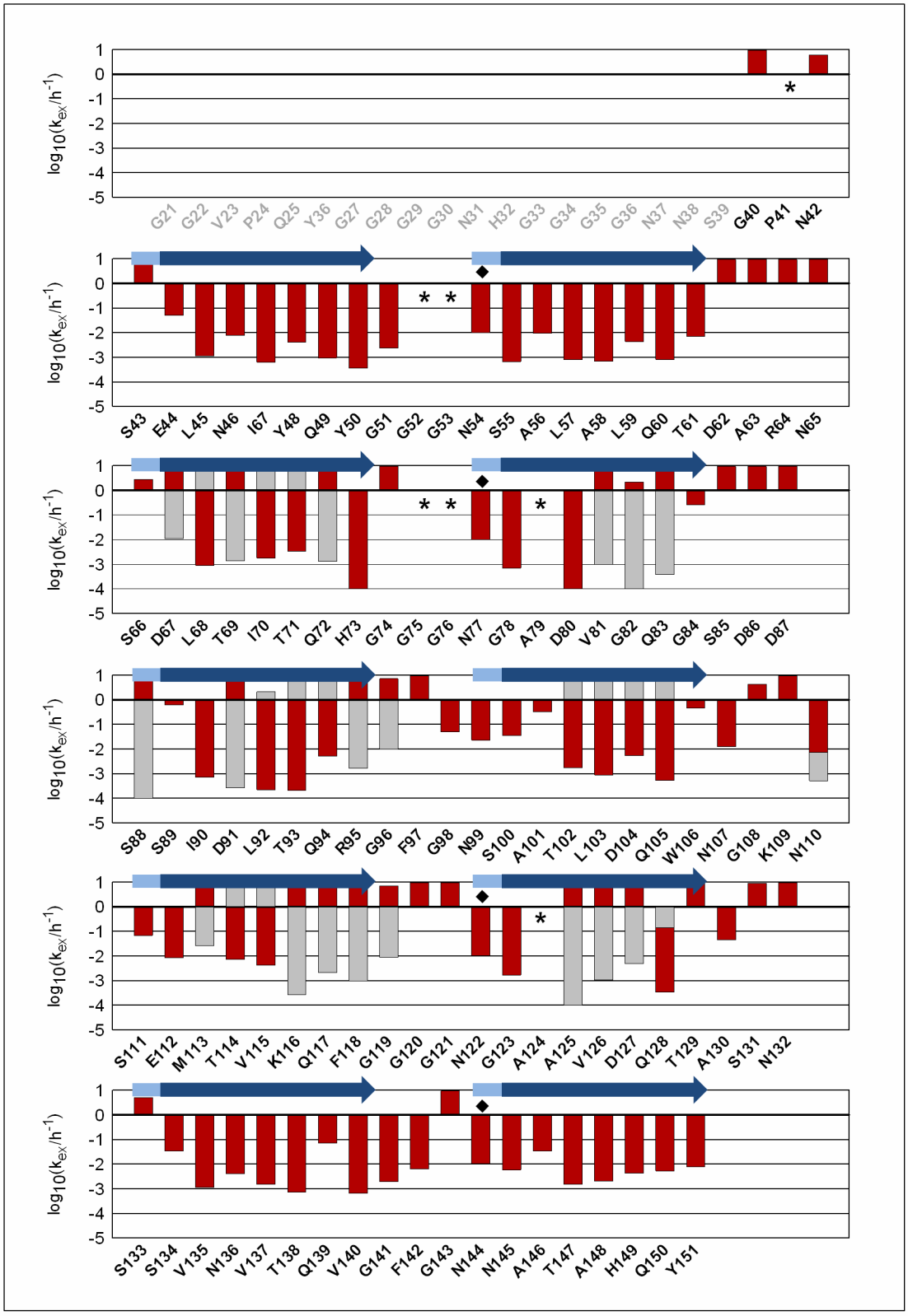

Figure App-6-3: H/D exchange NMR data of *Ec*CsgA$_{40\text{-}151}$ after seeding with *Ec*CsgB fibrils. Red and grey bars indicate exchange rates of *Ec*CsgA$_{40\text{-}151}$ k_{ex} (h^{-1}). Red bars correspond to the major population, grey bars to the minor population. Residues marked with a rectangle were analysed as one unit. Residues marked with an asterisk were excluded from analysis due to fast intrinsic exchange in DMSO (A79, A124) and missing assignments. Blue arrows indicate β-strands, which were localised in homonucleated *Ec*CsgA$_{40\text{-}151}$ fibrils.

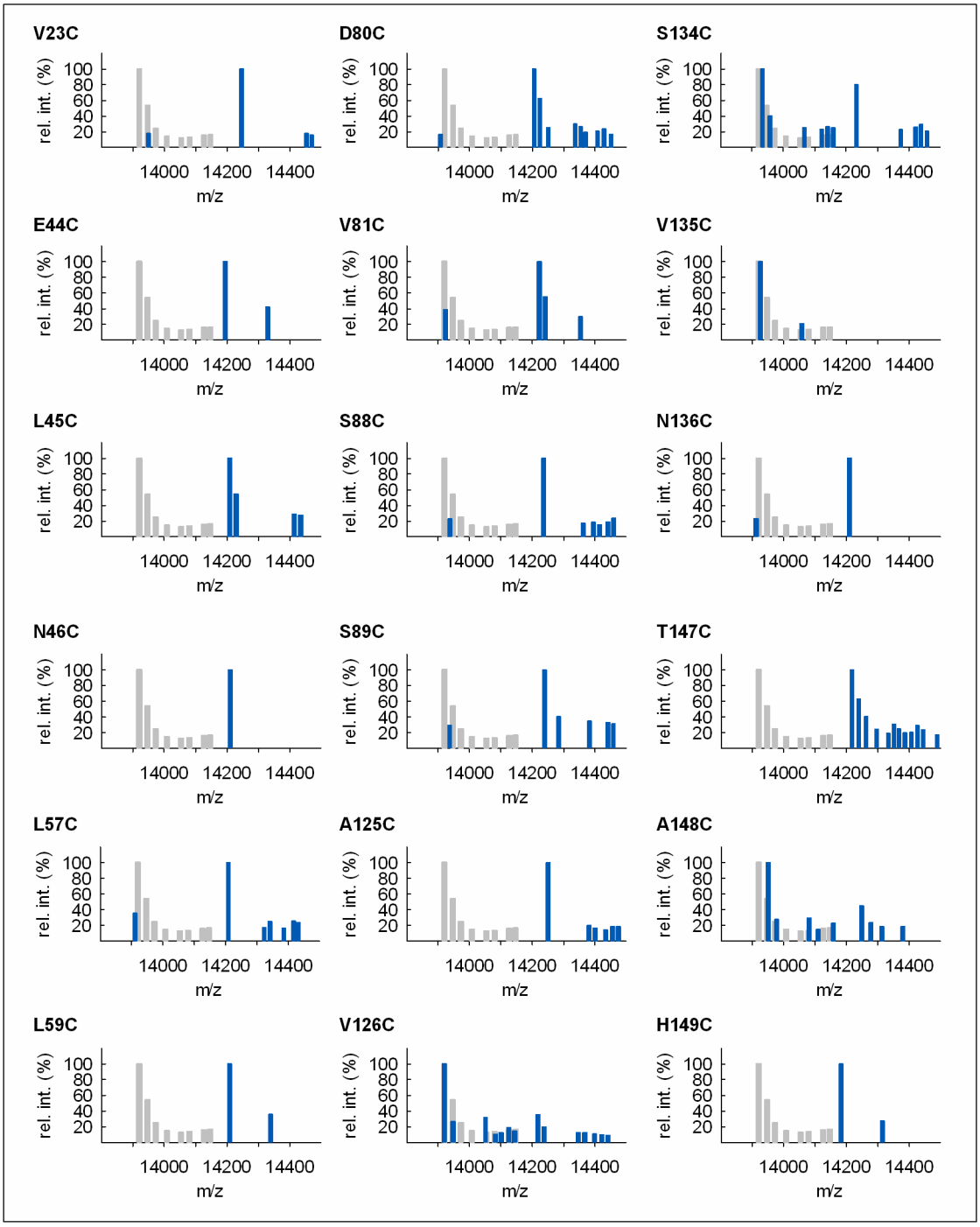

Figure App-6-4: MALDI spectra after thiol-labelling of *Ec*CsgA$_{21\text{-}151}$ cysteine variants with PM. Approximately 50 µg of wet fibrils were applied to mass analysis after PM-labelling. Covalent PM-labelling results in a mass shift of Δm = 297 Da. The observed masses were masked due to N-Formylation and a probably missing start methionine.

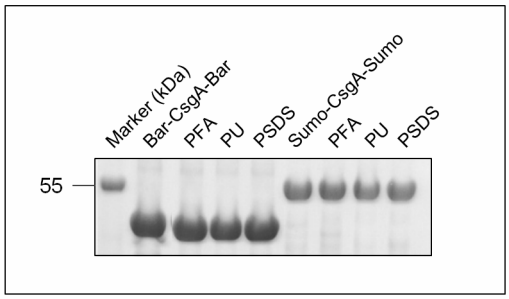

Figure App-6-5: Stability of aggregates formed by Bar-CsgA-Bar and Sumo-CsgA-Sumo. 4-12 % NuPAGE® (Coomassie staining). Pellet fractions were solubilised in 98 % (v/v) FA (PFA), 8 M urea (PU) and 0.1 % (w/v) SDS and loaded onto the gel after TCA precipitation.

D) List of Tables

E) List of Figures

Danksagung

Prof. Dr. Christiane Ritter danke ich herzlich für die Zeit in ihrer Arbeitsgruppe, die Möglichkeit zur Promotion im Bereich der Amyloidforschung und der damit verbundenen Herausforderung, ihre fachliche Unterstützung und die stete Förderung meiner Arbeit.

Prof. Dr. Michael Steinert danke ich für die bereitwillige Übernahme des Zweitgutachtens und seine Unterstützung im Rahmen meines Thesis Komitees.

Bei PD Dr. Michael Hust bedanke ich mich für die spontane und freundliche Übernahme des Vorsitzes der Prüfungskommission.

Prof. Dr. Dirk Heinz danke ich für die hervorragenden Arbeitsbedingungen in der Abteilung Molekulare Strukturbiologie, für seine Teilnahme an meinem Thesis Komitee und seine stete Diskussionsbereitschaft.

Bei Prof. Dr. Beat Meier und Dr. Christian Wasmer sowie Prof. Dr. Sven Saupe bedanke ich mich für die hervorragende Zusammenarbeit hinsichtlich des HET-s Projektes.

Dr. Thorsten Lührs danke ich für seine Hilfe bei der Durchführung verschiedener NMR-Experimente und seine Diskussionsfreude.

Dr. Manfred Rohde sowie Dr. Heinrich Lühnsdorf und ihren Kollegen gilt mein Dank für die Durchführung elektronenmikroskopischer Arbeiten und die Unterstützung beim Beladen unzähliger Grids.

Dr. Manfred Nimtz und seinem Team, vor allem Anja Meier und Undine Felgenträger, danke ich für die Aufnahme zahlreicher Massenspektren.

Rita Getzlaff und Beate Jaschok-Kentner danke ich für die Durchführung N-terminaler Sequenzierungen.

Für den regen Gedankenaustausch rund um die Curli Proteine danke ich allen Mitarbeitern der Arbeitsgruppe Makromolekulare Interaktionen. Madhu Nagaraj – you are next! Thank you for all the discussions about curli, fibril formation and H/D exchange NMR and the cookies in between.

Bei meinen ehemaligen Praktikanten Debora Reichmann, Janine Strehmel, Jonathan Crawford, Steffen Berthold und Torsten Steinbach bedanke ich mich für ihr Interesse an meiner Arbeit und ihre Unterstützung während verschiedener Etappen dieser Arbeit.

Mein herzlicher Dank gilt allen ehemaligen und derzeitigen Mitarbeitern der Abteilung Molekulare Strukturbiologie für die angenehme Arbeitsatmosphäre und ihre immerwährende Diskussions- und Hilfsbereitschaft – seien es die hervorragende Laborорganisation und technische Assistenz, die Unterstützung von Dr. Joop van den Heuvel bei Klonierungsfragen, die Hilfe von Dr. Joachim Reichelt bei Computerproblemen jeglicher Art oder die inspirierenden Kaffeepausen.

Besonders Susanne Freund, Ulrich Wiesand, Matthias Haffke, Christian Strube, Lilia Polle, Jennifer Block, Alexander und Sonja Eberth, Stefanie Loss, Carolin Schaper, Jens de Groot, Thomas Heidler, Ute Widow, allen voran den Laube-Aktivisten, Kletterfreunden und der Reisegruppe Sonnenschein um Luisa Winkler, Sabine Ernst und Thomas Krieg danke ich für die tolle Zeit im Labor und außerhalb dessen.

Ein großes Dankeschön geht an „meine“ drei Rudervereine Hallescher RC, Weißenfelser RV und Wolfsburger RC, die trotz meiner akuten Trainingsermangelung immer einen Rollsitz für mich frei hatten.

Dr. Victor Wray, Wiebke Zander und Andreas Fischer danke ich für das Korrekturlesen der Arbeit und der damit verbundenen Unterstützung auf den letzten Metern.

Das größte Dankeschön gilt meiner Familie. Besonders meinen lieben Eltern danke ich an dieser Stelle für ihre Unterstützung und ihren Rückhalt. Darauf immer und überall zählen zu können, macht vieles erst möglich.

Lebenslauf

Persönliche Angaben

Agnes Zimmer
geb. am 02.10.1982 in Weißenfels
ledig
deutsch

Schulische Ausbildung

1989 – 1991	Oberschule Valentina Tereschkova, Halle/Saale
1991 – 1993	Grundschule Brüder Grimm, Halle/Saale
1993 – 1996	Frieden Gymnasium, Halle/Saale
1996 – 2002	Goethe Gymnasium, Weißenfels Abschluss: Abitur

Akademischer Werdegang

2002 – 2007	Studium der Biochemie an der Martin-Luther-Universität Halle-Wittenberg Studienabschluss: Diplom (Note: gut)
2006 – 2007	Diplomarbeit an der Martin-Luther-Universität Halle-Wittenberg „Strukturelle und kinetische Charakterisierung der Oxalyl-Coenzym A Decarboxylase aus zwei verschiedenen Organismen“
2007 – 2011	Promotion am Helmholtz Zentrum für Infektionsforschung in Braunschweig in der Arbeitsgruppe: Makromolekulare Interaktionen, Abteilung: Molekulare Strukturbiologie

Publikationen/Tagungsbeiträge

Werther, T., Zimmer, A., Wille, G., Weiss, M. S., Golbik, R., Hübner, G., and König, S. (2010) New insights into structure function relationships of oxalyl CoA decarboxylase from *E. coli*. FEBS J. *10*

Werther, T., Zimmer, A., Wille, G., Weiss, M. S., Golbik, R., Hübner, G., and König, S. (2008) Kinetic and structural characterization of the oxalyl CoA decarboxylase from *E. coli*. Vortrag – 11th HEC Meeting, Greifswald